

Examining Listener Reaction Time as a Measure of Speech Disorder

R. A. Risi

University of Canterbury

2018

ABSTRACT

Historically, speech intelligibility has been the key measure of dysarthria severity, and both rating scales and orthographic transcription have been used as the primary forms of assessment. However, due to limitations in these measures, such as ceiling effects in transcriptions and subjective biases in rating scales, different forms of assessment are required to accurately assess speech severity in people with dysarthria. This study seeks to address these limitations by examining a new form of dysarthria assessment: listener reaction time. Listeners (one group of 16 younger adults, one of 16 older adults) performed three measures of speech intelligibility: a sentence transcription task, a rating scale task and reaction time task to veracity statements. These three measures were significantly correlated to one another, with reaction time accounting for 88% of the variance in transcription scores and 84% of the variance in the ratings of dysarthria provided by the younger listener group. Overall, the younger listeners produced significantly different reaction times when listening to speakers with dysarthria as compared to healthy control speakers. Reaction times between pairs of speakers (one with dysarthria, one control) were significantly different for all but the two speakers with the mildest dysarthria. In examining the older listener data, similar associations between assessment scores were found. However, the older listeners produced significantly longer reaction times than the younger group, and age and hearing loss were identified as significant factors for both reaction time and orthographic transcription.

TABLE OF CONTENTS

ABSTRACT.....	II
LIST OF TABLES.....	VI
LIST OF FIGURES.....	VII
CHAPTER ONE:.....	1
<i>INTRODUCTION.....</i>	1
1.1 Dysarthria	1
1.2 Listener Processing of Speech	2
1.2.1 Perception of Dysarthria	2
1.2.2 Evidence of Increased Processing Effort	3
1.3 Measuring Speech Intelligibility	4
1.3.1 Measurements from Listeners' Orthographic Transcription.....	4
1.3.2 Listener Rating Scale Measurements.....	6
1.4 Assessing Reaction Time to Auditory Stimuli.....	8
1.5 Veracity Statements.....	10
1.6 Effect of Variables	11
1.6.1 Listener Differences	11
1.6.2 Motoric Changes Associated with Age.....	12
1.6.3 The Effect of Age on Listener Processing	12
1.6.4 The Effect of Hearing loss on Listener Processing.....	14
1.7 Conclusions and Study Rationale.....	15
1.8 Aims and Hypotheses	15
CHAPTER TWO:.....	17
<i>METHODS.....</i>	17
2.1. Design.....	17
2.2 Participants	17
2.2.1 Speakers	17
2.2.2. Listeners	18
2.2.3 Recruitment.....	18
2.2.4 Exclusion Criteria.....	19
2.2.5 Demographics	19
2.3 Stimuli.....	20
2.3.1 Orthographic Transcription Stimuli	20
2.3.2 Rating Scale Stimuli.....	20
2.3.3 Veracity Statements	20
2.4 Equipment	21
2.5 Procedure	21
2.5.1 Pre-Assessments.....	21
2.5.2 Speech Assessments – Set up and Practice Trial	22
2.5.3 Orthographic Transcription - Sentence Intelligibility Test	23
2.5.4 Rating Scale	23
2.5.5 Reaction Time Experiment	24
2.5.6 Experiment Configuration.....	24
2.6. Data Treatment.....	25
2.6.1 Reaction Time - Data Treatment.....	25
2.6.2 Reaction Time - Excluded Data	26
2.6.3 Orthographic Transcription – Data Treatment.....	27
2.6.4 Rating Scale – Data Treatment	27

CHAPTER THREE:	29
RESULTS	29
3.1.1 Reaction Time Data - Plotting and Tests of Normality	29
3.1.2 Orthographic Transcription Data - Plotting and Tests of Normality	31
3.1.3 Rating Scale Data - Plotting and Tests of Normality	32
3.2 Research Question One – Difference in Younger Listeners’ Reaction Time to Dysarthric and Control Speakers	34
3.3 Research Question Two – Comparison of Reaction Time to Standard Measures	34
3.3.1 Orthographic Transcription Analysis	35
3.3.2 Rating Scale Analysis	36
3.4 Comparisons Between Measures	36
3.4.1 Reaction Time vs Orthographic Transcription	37
3.4.2 Reaction Time vs Rating Score	37
3.4.3 Rating Score vs. Orthographic Transcription	38
3.5 Research Question Three - Effect of Age on Results	38
3.5.1 Reaction Time Comparison	39
3.5.2 Effect of Age and Hearing Loss on Reaction Time	40
3.5.3 Effect of Age and Hearing Loss on Orthographic Transcription	41
CHAPTER FOUR:	43
DISCUSSION	43
4.1 Overview of Results	43
4.1.1 Difference in Reaction Time to Speakers with Dysarthria and Control Speakers	43
4.1.2 Reaction Time Compared to Standard Measures in Younger Listeners	45
4.1.3 Effect of Age on Findings	45
4.2 Study Limitations	47
4.2.1 Veracity Statements – Stimuli and Data Collection	48
4.2.2 Scoring of Veracity Statements	48
4.2.3 Limited Data Collection	49
4.2.4 Summary of Limitations	50
4.3. Summary of Findings	50
4.4 Future Research Directions	51
4.5 Clinical Implications	52
References	54
Appendices	60
Appendix A	60
Consent from University of Canterbury Human Ethics Committee	60
Appendix B	62
Speaker Demographic and Etiological Information	62
Appendix C	63
Listener Demographic and Pre Test Information	63
Appendix D	64
Sentence Intelligibility Test Sentences by Speaker	64
Appendix E	67
Veracity Statements and Ambiguity Ratings	67
Appendix F	73
Study Information Sheet for Listeners	73
Appendix G	76
Consent Form for Listeners	76
Appendix H	77
Demographic Questionnaire for listeners	77
Appendix I	78

Information for Listeners Requiring Full Diagnostic Hearing Test.....	78
Appendix J	79
Montreal Cognitive Assessment, and consent to use on behalf of Dr Ziad Nasreddine	79

LIST OF TABLES

<i>Table 1.</i> Demographic Averages and Pre-test Mean Scores by Group.....	20
<i>Table 2.</i> Speaker Demographic and Etiological Information	62
<i>Table 3.</i> Control Speaker Information	62
<i>Table 4.</i> Listener Demographic and Pre Test Information	63
<i>Table 5.</i> Veracity Statements and Ambiguity Ratings.....	67

LIST OF FIGURES

<i>Figure 1.</i> Bar chart frequency distribution of different reaction times to all speakers for younger listeners.	29
<i>Figure 2.</i> Mean (and interquartile ranges and extremes) reaction time for listeners to speakers with dysarthria and control speakers. Outliers identified with blue circles and red squares. ...	30
<i>Figure 3.</i> Bar chart frequency distribution of different reaction times to all speakers for older listeners.	30
<i>Figure 4.</i> Box and whisker plot of mean (and interquartile ranges and extremes) reaction time for older listeners to speakers with dysarthria and control speakers. Outliers identified with blue (control) circles and red (dysarthria) squares.	31
<i>Figure 5.</i> Bar chart frequency distribution of orthographic transcription scores to all speakers with dysarthria for younger listeners.	31
<i>Figure 6.</i> Bar chart frequency distribution of orthographic transcription to all speakers for older listeners.	32
<i>Figure 7.</i> Bar chart frequency distribution of orthographic transcription scores to all speakers with dysarthria for younger listeners.	33
<i>Figure 8.</i> Bar chart frequency distribution of different reaction times to all speakers for older listeners.	33
<i>Figure 9.</i> Median reaction time and 95 % confidence intervals displayed for control speakers and speakers with dysarthria, by younger listeners.	34
<i>Figure 10.</i> Median reaction time and 95% confidence intervals for individual speakers with dysarthria (red square) compared to their control speaker (blue circle) for younger listeners.	35
<i>Figure 11.</i> Median percent correct for orthographic transcription (1.0 = 100%, 0.5 = 50%) and 95% confidence intervals for individual speakers with dysarthria for younger listeners.	35
<i>Figure 12.</i> Median score and 95% confidence intervals for individual speakers with dysarthria by younger listeners. 0 = less intelligible, 1 = highly intelligible.	36
<i>Figure 13.</i> Percent correctly transcribed compared to average reaction time by speaker. Line of best-fit and dashed line representing 95% confidence interval for individual speakers.	37
<i>Figure 14.</i> Average rating score compared to the average reaction time by speaker. Line of best-fit and dashed line representing 95% confidence interval marked.	37
<i>Figure 15.</i> Percent correctly transcribed compared to rating score by speaker. Line of best-fit and dashed line representing 95% confidence interval.	38

Figure 16. Median and 95% confidence interval displayed for reaction time by group to both control speakers (blue) and speakers with dysarthria (red).39

Figure 17. Median and 95% confidence interval displayed for reaction time to individual speakers with dysarthria between the older and younger groups.39

Figure 18. Linear regression of reaction time by age of listeners in the older group. 95% confidence intervals displayed in the blue shaded area.40

Figure 19. Regression of effect of pure tone average on reaction time. Control speakers on the left (dysarthria = c) and speakers with dysarthria on the right (dysarthria = d).....41

Figure 20. Regression of effect of pure tone average on percent correct controlling for age. age 68-71 n =seven, age 72 -75 n= two, 76 -78 n= four, 79 – 82 n = one and 83 n = two.....42

LIST OF ABBREVIATIONS

dB	Decibels
dB HL	Decibel Hearing Level
Hz	Hertz
kHz	Kilohertz
MoCA	Montreal Cognitive Assessment
PTA	Pure-Tone Average
SIT	Speech Intelligibility Test
SNR	Signal to Noise Ratio

CHAPTER ONE:

INTRODUCTION

Dysarthria is a speech disorder commonly observed in people with damage to the central or peripheral nervous system (Darley, Aronson, & Brown, 1975). Historically, speech intelligibility has been the key measure of dysarthria severity, and both rating scales and orthographic transcription have been used as the primary tools of assessment (Weismer & Laures, 2002; Tikofsky, & Tikofsky, 1964). However, there are some issues with both forms of intelligibility measurement. Floor and ceiling effects, listener bias, and use of, or lack of, referent stimuli can all affect a clinician's ability to assess and monitor the severity of dysarthria (Hustad, 2006; Duffy, 2006; Weismer & Laures, 2002; Sussman & Tjaden, 2012). As a result, it may be difficult to identify the effects of mild dysarthria on speech processing or listener effort, and to reliably monitor changes in the same speaker over time. This study seeks to address these limitations by examining a new form of dysarthria assessment: listener reaction time. If it can be demonstrated that listeners take longer to process speech that is more severely affected by dysarthria compared to healthy speakers, this may become a useful adjunct in the quantitative assessment of the disorder.

1.1 Dysarthria

Dysarthria is a disorder that affects the efficiency and intelligibility of a person's speech. Darley & Brown (1969) described dysarthria as "a collective name for a group of speech disorders resulting from disturbances in muscular control over the speech mechanism due to damage of the central or peripheral nervous system" (p. 2). It is a motor speech disorder and can be chronic, transient, or progressively degenerative (Darley & Brown, 1969). There are many conditions known to cause dysarthria, and while in many cases the

underlying pathology is known before dysarthria occurs, in some diseases dysarthria is an early indicator of a neurological issue (Duffy, 2006).

Dysarthric speech can be evaluated in a number of ways. Informal perceptual observations are common in clinical practice, where key descriptors are used to describe the speech of an individual. Words like “slurred,” “jerky,” “breathy,” and “laboured” are all examples of possible descriptors, as reported by Darley et. al (1975). Some dimensions of speech disorder can be also examined using objective assessments, for aspects like reading rate, vowel prolongation, or physiological measurements of hyper-nasality (Kreul, 1972). However, to capture the effect of dysarthria on everyday communication, measurements that explore the listener’s ability to process and comprehend the speech signal are necessary.

1.2 Listener Processing of Speech

There are several cognitive processes required for the perception and understanding of speech. These processes allow the listener to rapidly match a speaker’s utterances against a large number of possible lexical items, recognize when a break between words occurs, segment the speech stream into these individual words, and assign the correct word to each of these segments. The meaning of these words is then determined, and the individual words are comprehended within the context of the sentence (McAuliffe, M. J., Borrie, S. A., Good, P. V., & Hughes, L., 2010).

1.2.1 Perception of Dysarthria

When listeners hear disordered speech it creates phonemic uncertainty, as a larger number of possible lexical targets have to be considered (Liss 2007). Multiple processes are thought to work together to make sense of disordered speech, beyond what is required for listening to clear speech.

To many people, a speech disorder is considered to be only a deficit in the production of speech. In research, however, is accepted that there is a bidirectional aspect to the disorder,

occurring between the speaker and listener, that affects communication competency (Borrie, McAuliffe, Liss, Kirk, O'Beirne, & Anderson, 2012). In addition to the articulatory impairment that causes phonemic ambiguity, Liss (2007) hypothesized that prosodic and voicing impairments experienced by speakers with dysarthria might produce further challenges for listeners in comprehending words.

This hypothesis is based on research by Liss, Spitzer, Caviness, Adler, and Edwards (1998) and Liss, Spitzer, Caviness, Adler, & Edwards, (2000). In 1998, Liss et al. looked at speech samples from 70 speakers with hypokinetic dysarthria and found that the listeners' ability to comprehend the disordered speech signal was reduced due to the monotonous nature of their speech pattern. A more pronounced negative effect on intelligibility was found in Liss et al.'s 2000 study with ataxic dysarthria. This suggests that differences in intelligibility exist between different forms of dysarthria. Overall, dysarthria is related to an increase in difficulty in speech perception for listeners (Liss, 2007).

1.2.2 Evidence of Increased Processing Effort

Increased processing demands associated with dysarthria might occur very early upon hearing a word or phrase. Theys and McAuliffe (2017) had listeners hear speech samples from speakers with dysarthria and control speakers and measured two neurological waveform responses using electroencephalography, the N100 (a measure of earlier sensory processing), and the N400 (a measure of later cognitive-linguistic processing). The N100 showed increased amplitude and decreased latency for dysarthric speech compared to the control speakers, while no significant differences were found in the N400. Theys and McAuliffe postulated that the quality of dysarthric speech affects early sensory processing. This additional attention to early differences in the speech signal primes listeners to allocate more resources to decoding the signal than what is required for clear speech.

1.3 Measuring Speech Intelligibility

Auditory-perceptual assessment methods are the current “gold-standard” for diagnosing and monitoring overall dysarthria severity (Bunton, Kent, Duffy, Rosenbek, & Kent, 2007). There are two main ways in which intelligibility can be measured: transcription (numeric count of words understood), and subjective rating. Values from both assessments are typically used to index how easy it is for listeners to understand speech. These two assessment methods will be discussed in the following sections.

1.3.1 Measurements from Listeners’ Orthographic Transcription

Historically, speech intelligibility has been the main focus in the assessment of communication and dysarthria severity. Speech intelligibility is traditionally measured by calculating the percentage of words a listener correctly identifies in a sentence, or the percentage of phonemes correctly identified in a word. A higher percentage of errors is considered to reflect less intelligible speech (Yorkston & Beukelman 1981). This form of assessment is often referred to as orthographic transcription. This is commonly used in standard clinical intelligibility tests, where a listener will hear a speech sample from a person with dysarthria, then repeat or type what they heard (Yorkston, Beukelman, & Tice, 1996). The original transcript is compared to the words identified by the listener, to see how often and what kind of errors are made. In their seminal paper, Tikofsky & Tikofsky (1964) used these procedures to demonstrate a significant difference in the intelligibility of speakers with dysarthria compared to those without. Differences were also found between different speakers with dysarthria, which led to further work looking at orthographic transcription as a measure of classifying dysarthria severity (Tikofsky, 1970; Hustad, 2008). Performance on this measure is easily understood in terms of its effect on communication, with someone who is less intelligible being harder to understand (Yorkston & Beukelman, 1981). There is also typically very high inter-rater reliability with sentence intelligibility measures, with dos

Santos Barreto, & Ortiz (2015) finding that there was highly significant inter-rater reliability in the transcription method undertaken at $r = .94$, from the 120 listeners studied.

However, this system of measurement also has its flaws, as it is not sensitive to all degrees of dysarthria impairment. This due to its susceptibility to floor and ceiling effects. For example, a speaker may have speech that is clearly affected by dysarthria (e.g. speech that is slurred, breathy, with an unnatural rate or rhythm) but still have 100% of their words correctly identified by listeners. In contrast, some speakers who are unable to produce any intelligible words can continue to experience a decline in their ability to make speech movements which cannot be reflected in transcription based assessment. This test is also not entirely ecologically valid, as functional communication requires listeners to extract meaning from long strings of words, not simply remember, identify and repeat words.

A study by Sussman & Tjaden (2012) compared orthographic transcription with rating scale estimates, for healthy control speakers, speakers with multiple sclerosis and speakers with Parkinson's disease. Listeners were required to rate speech samples on a 0-1.0 scale, from 0-no impairment to 1.0- severely impaired based on their overall impression of the speaker. The scale estimate of intelligibility was more sensitive to speech impairment than the word or sentence intelligibility calculated from their transcriptions. This finding suggests that orthographic transcription measurements may not be able to index subtle changes to the quality of an individual's speech.

These limitations were also discussed by both Hustad (2006) and Patel, Usher, Kember, Russell & Laures-Gore (2014). Patel et al (2014) inferred that orthographic speech intelligibility assessments are influenced by speaker and listener variables, which make these values sensitive to small changes in the testing procedure. Therefore, it is possible to generate different results based on the test procedure employed. They also discussed the role that context and familiarity can play in comprehending dysarthric speech, citing previous studies

like Hustad, (2007) and Yorkston & Beukelman, (1981) who found that speakers receive a higher intelligibility score when using sentences instead of single words, due to the added context.

Sentence intelligibility tests can also be time-consuming. To ensure accuracy, these tests need to be repeated over a number of trials which can then be averaged. This can involve listeners hearing and transcribing at least 22 sentences (Yorkston et al., 1996) to get an overall percent correct for each speaker. A single sentence is not sufficient in replicating this average correct response.

1.3.2 Listener Rating Scale Measurements

In clinical research, rating scales are often used to quantify listeners' perceptual impressions of speech, such as a speaker's level of naturalness, intelligibility, or speech severity. To assess intelligibility, listeners are commonly asked to either rate the speakers' intelligibility or rate "how easy" the speaker is to understand (Kim et al., 2011; Tjaden & Wilding, 2004; Turner et al., 1995; Weismer et al., 2001). For example, Anand & Stepp (2015) used a visual sort and rank method, where participants compared and rated speech samples on a continuum, from most to least intelligible. Weismer & Lares (2002) utilised a reference stimuli against which different speech samples could be numerically compared.

There are some limitations inherent to these rating methods. For example, both Anand & Stepp (2015) and Weismer & Lares (2002) relied on reference stimuli to accurately compare different samples of dysarthria. This can be problematic considering different referents can change listeners' ratings of dysarthria. Weismer & Lares (2002) investigated the effect of choosing different referents and found that ratings did change when different speech samples were used as the referent.

In many other rating methods, no referent is used, and listeners are simply provided with an ordinal scale, (from 0- 10 or from 0-1.0) to rate speech samples. As discussed in the opinion piece by Bradburn and Miles (1979), these ordinal scales provide limited information about dysarthria, as they do not enable people to rate fine-grained differences between speakers. Another issue with ordinal rating scales is the fact that the magnitude difference between numbers on these scales are often not equal, which can lead to issues when attempting to average scores across listeners. For example, the amount of change from a score of four to a score of five may be different than the amount of change from a score of one to two. In this way, a rating scale may not be linear, and is not an interval measure, although it is often discussed as though it were. This is another significant issue with this type of assessment.

In any type of listener rating procedure, inherent subjectivity also produces meaningless variation in scores. Thus, it can be difficult to compare and interpret rating scale scores that are provided by different listeners. In addition, it has been hypothesized that listeners might express some biases in their ratings of intelligibility, which are unrelated to the comprehensibility of the words spoken (Fletcher, McAuliffe, Lansford, Sinex, & Liss, 2017). For example, Street, Brady & Putman, (1983) found that listeners exhibited a natural bias toward speakers with rates of speech similar or slightly faster than their own speech rate. This is likely to affect their ratings of intelligibility in speakers with mild dysarthria. There is no clear way to circumvent this natural bias, but it must be acknowledged as a flaw in relying on rating scales to index information about auditory stimuli. Ultimately, rating scale scores cannot be translated to a functional measurement of communicative success. Merely identifying that someone appears more or less intelligible than a previous sample, or assigning a numerical value to their intelligibility, cannot directly address the listener's ability to process the speech sample.

Clearly, there are issues with using both rating scales and orthographic transcriptions in the assessment of dysarthria. This has led researchers and clinicians to consider other methods of assessment that may ameliorate some of these limitations. One possible solution is using listener reaction time.

1.4 Assessing Reaction Time to Auditory Stimuli

The idea of using reaction time as a speech assessment tool is not new, and it is often used as an objective measure to perceptual stimuli that can otherwise be difficult to assess. It has been used previously in other disciplines with physical and visual stimuli; Hyman (1953) discussed how reaction time to visual stimuli can vary if there is a greater number of stimuli to choose from, and if the probability of the target stimuli varies.

Reaction time has the advantage of being an objective measure, as the physical action of the response can be recorded and repeated, and there is no dependence on listener judgement of the speaker. A broad range of stimuli can be used, and this assessment can be adjusted depending on the research question it is addressing. It is understood that reaction time can be used as a measure of mental demand, with a longer reaction time indicating increased cognitive workload (Evitts et al 2016). Prolonged reaction time has been used numerous times as evidence of increased cognitive difficulty and processing time. (Gough 1965; Munro & Derwing, 1995; Gatehouse & Gordon, 1990; Houben, van Doorn-Bierman & Dreshler, 2013). It can be a useful measure for comparisons between groups, or with the same speakers over time. This makes it an ideal form of measurement when assessing dysarthria.

In terms of auditory perception, reaction time has been used to understand how quickly people are able to comprehend different aspects of language (Gough 1965). More recently it has been used to understand differences when listening to foreign vs familiar

accents (Munro & Derwing, 1995; Wilson & Spaulding, 2010), synthetic speech vs live speech (Reynolds and Fucci 1998; Evitts & Searl 2006), and speakers with and without dysphonia (Evitts et al 2016).

Across studies, listeners' reaction times have consistently been higher when exposed to more challenging listening conditions. For example, Munro and Derwing (1995), compared reaction times for listeners hearing native English speakers and Mandarin-accented English speakers. They found that despite both groups of speakers often being fully intelligible to listeners, listeners rated the accented speakers as more difficult to understand, and took longer to decipher the message. In this way, orthographic transcription did not reflect the findings of a rating scale as well as the reaction time measure, due to lack of sensitivity to mild differences in speech. They hypothesized that this was a reflection of the greater difficulty they had with processing the message. This is due to listeners first needing to replay the message from short-term memory to comprehend what they heard before they could act on it. This would require significantly more cognitive resources, which may be reflected in the fact that these speakers were rated as more difficult to understand, despite being completely intelligible.

Gatehouse & Gordon (1990) used reaction time as a measure of listening effort for listeners with and without the use of hearing aids, in quiet and in noise. They used reaction time to verify that hearing aids work to reduce the processing effort for listeners. They had people with hearing loss listen to speech in quiet and then speech in noise, with their hearing aids on and then off in both conditions. Listeners were required to repeat back what they heard from a sample sentence and were able to respond quickest with their hearing aids on in the speech alone condition. They found that reaction time was able to verify the advantage of hearing aids', and this measurement was more sensitive to hearing aid use than speech intelligibility alone, as the intelligibility measure was limited by ceiling effects. Even when

listeners could still correctly repeat 100% of the sample sentence, the speed at which they could do this differed based on the presence of background noise and the use of hearing aids.

Reaction time was also the measurement used by Houben, van Doorn-Bierman & Dreshler (2013) for their study looking at listening effort. Listeners heard a series of numbers and had to either add digits together or repeat the final digit they heard. This was done in varying signal to noise ratios (SNRs). They found that for both tasks participants responded quicker in better (higher) SNR. This was true regardless of how accurate participants were; even when they were able to accurately hear and repeat the correct answer, it still took longer in a more complex listening situation.

A variety of stimuli can be used in a reaction time experiment. When assessing reaction time to speakers with dysphonia, Evitts et al (2016) had listeners indicate whether a word that was visually presented was the same or different as one they were hearing. Gatehouse and Gordon (1990) had listeners repeat back what they heard from a sentence, and also used speech shaped stimuli where participants had to push a button to indicate when they could hear something that sounded like speech. Houben et al (2013) used digits as their stimuli, with participants adding two digits together or responding by repeating the final digit they heard, for comparisons of tasks of differing complexity. Munro and Derwing (1995) in their comparison of accented vs native English speech used veracity statements, as did Wilson and Spaulding (2010).

1.5 Veracity Statements

Veracity statements have been used successfully in previous research, and they can be used as a consistent measure of reaction time. A veracity statement is a true or false statement that is straightforward and can be easily determined using everyday knowledge. Listeners are required to listen to a sentence and then respond (select true or false) based on the

information they heard. There are several reasons veracity statements make sense as the stimuli for a reaction time measurement of speech intelligibility. Compared to other reaction time stimuli, veracity statements offer a direct real life comparison that digit recall or sentence repetition cannot offer. Listening to a sentence and responding based on the information heard is, in essence, the very nature of communication. Using this measure, that more closely resembles everyday life, allows for a better comparison to real-world ability.

Manous, Pisoni, Dedina & Howard (1986) used veracity statements to investigate reaction time to natural voices and synthetic voices. They had true and false sentences that were either three or six words long. They found significant effects of type of voice on reaction time, and also found that word length had an effect on the speed of verification, with the six-word sentences taking longer on average.

Munro & Derwing (1995) in their study comparing reaction time in response to native speakers of English and Mandarin used a word list of sentences between five and eight words long. Wilson and Spaulding (2010) in their comparison of native and foreign-accented speech used the same words lists as Munro & Derwing (1995). For both studies, reaction time was significantly different between accented and native speakers.

1.6 Effect of Variables

1.6.1 Listener Differences

While the benefits of using speech intelligibility as a primary measure of dysarthria severity have been demonstrated, a caution had been identified with regards to methodology. Schiavetti (1992) stated, “any measure of speech intelligibility is a measurement of the interaction between a speaker, a transmission system and a listener” (p.12). He further discussed how parameters of each of these three variables need to be quantified, and assessing speech intelligibility can be easily affected by changes any of these variables.

While the transmission system can be carefully controlled and consistently used, and speakers are typically the variable of interest, there are many unknown variables inherent to different listeners that could affect the reliability of the results in speech intelligibility measures.

A significant issue in any measurement used for diagnosing dysarthria severity is discrepancies in the control of these listener variables. Several of these variables are discussed below, but not all variables can be controlled for or assessed in a psychoacoustic experiment.

1.6.2 Motoric Changes Associated with Age

It is widely acknowledged that many motoric movements slow with age, with older adults typically demonstrating longer movement times than younger adults (Lyons, Elliott, Swanson, & Chua, 1996). This is due to physiologic components like muscle strength, which has been reported to decline from middle age onward (Bassey & Harries, 1993; Frontera et al., 2000). Notably, the decline in muscle strength associated with aging is greater than the decline observed in someone with a reduction in muscle mass (Kallman, Plato, & Tobin, 1990).

As differences in motoric movement times are not a key focus of this research study, this anticipated difference in movement speed should be controlled in order to assess the true reaction time taken to perceptually process the speech signal.

1.6.3 The Effect of Age on Listener Processing

It is not just physical changes that occur with age. Age-related cognitive changes are considered to be a contributing factor to processing difficulties in the elderly population (Committee on Hearing, Bio-acoustics, and Biomechanics [CHABA], 1988). This means that when assessing processing time to degraded speech signals, it is necessary to consider whether age alone is creating a difference in listeners' ability to react to speech samples,

compared to younger listeners. A consistent finding with studies utilizing reaction time measures is that older listeners have a slower response time compared to younger counterparts (Salthouse, 1985). This is true especially if the task becomes more complex, and cognitive demand increases (Birren, Woods, & Williams, 1979; Cerella, Poon, & Williams, 1980). This is largely attributed to central processing changes that occur with age. Bashore, Osman & Heffley in 1989 (who conducted a meta-analysis of studies looking at reaction time and P300 responses) found that although the rate of mental processing speed is generally slower among older persons, there are some aspects of cognition which are more sensitive to the increase in processing demands demonstrated by longer reaction time tasks. In contrast, there are other aspects of cognition that remain insensitive to increased difficulties in mental processing, and reaction time studies do not highlight these elements. This may contribute to the variability in performance in processing speed assessments that still exists within the older population.

A meta-analysis conducted by Verhaeghen & Salthouse, (1997) examining the effects of age on a range of cognitive factors. They determined that under the age of 50, younger listeners process stimuli faster (youngest ages tested being 18 years), and that older listeners were exponentially slower.

Indeed, this finding was supported by Kangas & Allen, (1990), who investigated the ability of older listeners with and without hearing loss to understand normal and synthetic speech using transcriptions. They found that listeners with normal hearing were significantly more accurate in transcribing both normal and synthetic speech. The effect of hearing loss on speech comprehension has clearly been demonstrated in audiological literature, and this paper reinforced this finding across different types of speech (normal and synthetic). Kangas & Allen also observed that even in the normal hearing category, the listener who individually had the lowest score by a marked amount was the oldest listener, who was 68 (mean age for

normal listeners was 56 years). Hence, they postulated that age alone may have contributed to a lower intelligibility score for this particular listener.

Age has been consistently cited as a factor affecting processing speed over and above the effects of hearing loss alone. For example, a study by Divenyi, Stark, and Haupt (2005) found that speech understanding declined more in people over 70 than what is expected from the rate of decline in their hearing thresholds alone. This suggests that while for mid-aged listeners (below 65 years approximately) with a hearing loss, the main barrier to speech processing is the effect of the hearing loss, for older adults (over 65 years), age-related changes to cognition may play the larger role.

1.6.4 The Effect of Hearing loss on Listener Processing

As stated above, the presence or absence of a hearing loss has a significant effect on listeners' ability to identify and comprehend words. While dysarthria can occur at any age, it often presents later in life. This means the communication partners of speakers with dysarthria are more likely to experience age-related hearing loss. Moscicki, Elkins, Baur, & McNamara (1985) reported that age is considered the biggest risk factor in developing a hearing loss, with approximately 83% of people over the age of 70 having a hearing loss. Weinstein (2014) reported hearing loss is greatest in people over 70 years of age, but it can begin to occur in people as young as 50, especially if they have other factors such as genetic predisposition or an environmental history of noise exposure contributing to the degeneration of important hearing cells.

As reported by Gates (2005): "(it is) useful to regard presbycusis (age-related hearing loss) as a mixture of acquired auditory stresses, trauma, and otological diseases superimposed upon an intrinsic, genetically controlled, aging process" (p. 1111). The degree and progression of hearing loss can vary, but it is often a barrier to understanding what others are saying. Even in an optimal environment, with a clear speaker, hearing impairment can

severely impact how much is understood. Listening to a speaker who has dysarthria adds a greater challenge, as the pathway to communication now has interference to both speech production and speech reception (McAuliffe, Borrie, Good, & Hughes, 2010).

1.7 Conclusions and Study Rationale

In summary, reaction time offers a valuable method of indexing the increased processing effort required to understand speech in challenging conditions, as well as quantifying the additional difficulties faced by older listeners due to age and hearing loss. Reaction time measures are objective, functionally interpretable, and can be completed relatively quickly. These features are ideal when trying to track and monitor a speech disorder over time, due to the varying nature of its progression. In addition, the apparent sensitivity of reaction time measurements makes them ideal for quantifying communication problems in the earlier stages of disease progression, where orthographic measures may fail.

Bashore et al, (1989) concluded that there is a consistent decline in mental processing speed among elderly adults. Assessing the individual effects of hearing loss and age on the processing of dysarthric speech will provide insight into the additional difficulty faced by communication partners when there is both hearing loss and dysarthria affecting communication. This is an area of research that is often neglected in studies of motor speech disorders. However, based on the prevalence of both dysarthria and hearing loss in the elderly population, it is an important aspect to consider in any new measurement of communicative disorder.

1.8 Aims and Hypotheses

This study has several aims. The primary purpose is to determine whether it is viable to use a reaction time measure as a sensitive test of dysarthria severity, and whether this measure is comparable to previously established forms of dysarthria assessment. This study

will also examine the impact of age and hearing loss on reaction time, comparing the results from younger listeners and older listeners and examining the effect of hearing loss in the older listeners. To address these aims, specific research questions have been created below:

1. Does reaction time differ when listeners are exposed to speakers with dysarthria compared to healthy speakers?
2. Are results from a reaction time paradigm similar to those obtained from common measures (orthographic transcription and rating scales)?
3. Is there an effect of age and hearing loss on listener-based measurements of dysarthria?

Based on available literature regarding perceptual studies of auditory information, the following hypotheses have been made with regard to the research questions.

1. Listener reaction time will be longer when responding to dysarthric speech compared to speech from healthy adults.
 - 1b. These effects will be strongest for speakers with more severe dysarthria.
2. Measurements of reaction time, orthographic transcription accuracy, and rating scale scores will be highly correlated.
3. Older listeners will have similar results to the younger listeners (in terms of comparisons between control speakers and dysarthria speakers), but will have overall slower reaction times.
 - 3b. Hearing loss will have a significant effect on scores.

CHAPTER TWO:

METHODS

The following chapter is about the methodology of the experiment used to investigate the research questions, including information on the recruitment and demographics of the listeners, stimuli, speakers, and procedures employed. It also outlines the procedure for statistical analysis of the generated data.

As per University of Canterbury policy, approval to undertake this study was sought from the University of Canterbury Human Ethics Committee. Approval was acquired on 3 July 2017, and again after an amendment was requested on 18 July 2017. Copies of both letters are available in Appendix A. The procedures and recruitment discussed below were conducted based on the information provided in these applications.

2.1. Design

This experimental study first utilized a two by two design, with age group of a listener (older or younger) and presence of dysarthria (dysarthria or control speaker) as the independent variables for the reaction time test. There was also a correlational aspect to this experiment, with three distinct forms of assessment. This study was non-blinded, as participants were aware of the overall purpose of the study.

2.2 Participants

2.2.1 Speakers

Male and female speakers of New Zealand English with a diagnosis of dysarthria between the ages 43 – 80 provided speech samples that were re-analysed in this study. There were initially 12 speakers recorded and after evaluating the quality and consistency of the recordings, a total of eight speakers were included. These speakers were considered to

represent wide a spectrum of mild through to severe dysarthria severity. Eight control speakers were also recruited who approximately matched the speakers with dysarthria for both age and gender. Demographics and etiological information about the included speakers can be found in Appendix B.

2.2.2. Listeners

To answer the research questions, two groups of listeners were required. Participants were grouped by age, with the younger group all between the ages of 18-28 years old and the older group between 65-83 years old, with 16 participants in each group.

Munro and Derwing (1995), enlisted 20 participants per listening group. For initial analyses, 16 per group (younger and older) was considered a reasonable number. Each listener provided 120 responses for the reaction time experiment, and 44 for the orthographic transcription task.

2.2.3 Recruitment

To participate in this study, younger listeners needed to be between 18-30 years old, and have hearing thresholds within normal limits (defined as hearing thresholds at or above 20 dB across four frequencies, .5 kHz, 1 kHz, 2 kHz, and 4 kHz). Older listeners needed to be between the ages of 65 – 90, and have normal hearing or hearing loss attributed to age-related changes to the hearing organ. Participants also needed to pass a cognitive screen to meet inclusion criteria (Nasreddine et al 2005). All listeners needed to be fluent speakers of New Zealand English, either being born and raised in New Zealand or having lived in New Zealand since childhood. Participants in the younger age group were recruited as part of their coursework for the paper CMDS161 at the University of Canterbury. They had the option of taking part in this study to receive three course credits. Other participants in the younger group were recruited from within the University of Canterbury and received a \$20 Motor Trade Association (MTA) voucher for reimbursement of their time.

The UC Speech Production- Perception Lab database of research volunteers was used to recruit participants for the older age group. This database contains the names and contact details of individuals who have participated in previous research studies and had indicated they were happy to be contacted again for further research studies. Prospective participants were contacted at random if they met the initial age range requirement, and were included based on their willingness to participate if they meet the eligibility criteria until adequate numbers were reached. Participants were all initially contacted via email, which included a copy of the information sheet. All participants in the older group received a \$20 Motor Trade Association (MTA) voucher as reimbursement of their time.

2.2.4 Exclusion Criteria

Exclusion criteria involved 1) being an age outside of the two test groups, 2) any co-occurring medical, physical or cognitive issues that would make participation in testing unreliable and 3) anyone who would be unfamiliar with listening to New Zealand English, as unfamiliarity with the language and accent is likely to add time to listeners' responses, and make their data significantly different to the other participants.

2.2.5 Demographics

Information about listeners can be found in the Appendix C. A total of 32 listeners were included in this study, 16 in each group. Each listener participated in all three measurements. Details of each group are included in Table 1 below, along with results of the pre test measures also included in this study that will be discussed further within the procedures section.

Table 1. Demographic and Pre-test Mean Scores by Group.

Group	Number	M age	Age range	M PTA	M MoCA	Sex
Younger	16	21.9	18-28	11.25	28.69	M=1, F=15
Older	16	74.2	69-83	31.56	27.44	M= 7, F= 9

note. M = mean, PTA = pure tone average of both ears, MoCA = Montreal Cognitive Assessment, score of 26 or greater required to pass.

2.3 Stimuli

2.3.1 Orthographic Transcription Stimuli

Each speaker was recorded saying 11 sentences ranging from five to fifteen words in length. These sentences were sourced from the Speech Intelligibility Test (SIT) Manual (Yorkston et al., 1996), and are displayed in Appendix D.

2.3.2 Rating Scale Stimuli

Each speaker also said the sentence, “He slowly takes a short walk in the open air each day” which is an excerpt taken from ‘The Grandfather Passage’ (Van Riper, 1963). This sentence was chosen as a comparison point as it is long enough to give listeners time to hear each speaker, and is from a passage consistently used in speech measurements.

2.3.3 Veracity Statements

True or false statements were taken from a list that was provided from an electroencephalography study of speech perception. Stimuli were then further selected with a preliminary check of ambiguity based on researcher selection. From the remaining sentences, each speaker was directed to read aloud 20 sentences, half of which were true and the other half false. After recording, these sentences were reviewed and some sentences were excluded that were deemed inappropriate to use (if participants laughed or coughed mid-recording).

The remaining sentences were then checked further for ambiguity, using a rating system by five independent raters, who selected whether a statement was true or false, and

how confident they were in their answer on a one to five scale (one meaning not at all confident, could have been either, five meaning very confident, cannot be the opposite answer). Any sentence that received a one or two was immediately excluded, all correctly identified sentences with a score of four or five were included. A table of the ambiguity test for all sentences is included in Appendix E. Of the original 20 sentences read aloud, 15 per speaker were used in this study.

Noise was removed from the speech samples gathered from the control speakers, using the default method of spectral subtraction in PRAAT. Each file was individually cut to contain only one statement per file, with the file beginning with the onset of the first word and ending after the final word was completely finished. This was determined by visual inspection of the waveform and repeated listening to the sentence to determine start and finish times.

2.4 Equipment

An Interacoustics AS608 Screening Audiometer with Peltor H7A headphones was used to test listeners' hearing prior to the experiments. The experiments were all conducted using MatLab software, on a Dell desktop computer at the University of Canterbury. Participants listened to the auditory stimuli using Sennheiser HD 201 headphones.

2.5 Procedure

2.5.1 Pre-Assessments

Participants who agreed to participate attended an appointment at the University of Canterbury. They reviewed the information about this study from the information form (Appendix F), and if they agreed to participate signed the consent form (Appendix G), and demographic questionnaire (Appendix H). They completed a pure tone hearing screen, which tested the main four speech frequencies (0.5 kHz, 1kHz, 2kHz, and 4 kHz.). No participants

in the younger age group were identified as having a hearing loss. In the older age group, participants with hearing loss that appeared to be predominately due to age-related hearing damage were still included in the study. The pure tone average (PTA) for each participant was identified by averaging the pure tone thresholds recorded at the four frequencies tested. Any participants with a significant hearing loss (over 20 dB at any frequency) were given the option to book in for a full diagnostic hearing test at the University of Canterbury free of charge in the volunteer clinic (Appendix I).

The Montreal Cognitive Assessment (MoCA©) was used with permission of Dr. Ziad Nasreddine, Neurologist, to screen cognitive function, as an internationally recognized screen of cognition (Nasreddine et al 2005). This contains components assessing visuospatial/executive function, naming, memory, attention, language, abstraction, delayed recall, and orientation (Appendix J). A score of 26 or higher is required to pass. The results for each participant in these pre-assessments can be found in Appendix C.

2.5.2 Speech Assessments – Set up and Practice Trial

Participants were seated in front of a desktop computer screen with a keyboard placed on the desk in front of them. Participants listened to the speech samples through the headphones plugged into the computer in this position.

The volume was set to 50 percent volume capacity on the desktop before each listener began and a practice trial was played, where participants could listen to speech samples from speakers with dysarthria (additional speakers who were not included in the current study). This practice trial included five sentences that were true or false, and participants responded by pushing a button to indicate their answer (N for true M for false). They could use any system that was comfortable to push the buttons, often this being either the index finger of each hand on a response button or the dominant index finger used to push both buttons. After hearing these sentences participants had the option of increasing or decreasing the volume for

the remainder of the computer tasks. All of the listeners in the younger group were happy to have the volume at the pre-set level, while six of the older listeners had it increased by one level, raising it to 60 percent capacity.

This practice trial also served as a primer for listeners to hear what a speaker with dysarthria can sound like, and to practice the task for the reaction time experiment. Speakers who were used in the practice trial were not used in any other task to prevent familiarity with speakers and sentences.

2.5.3 Orthographic Transcription - Sentence Intelligibility Test

Experiment One consisted of listeners hearing sentences from the SIT, read by four of the eight speakers with dysarthria, with 44 sentences played in total. Listeners transcribed each sentence individually by typing the words they could hear into a textbox as per the Yorkston et al. (1996) SIT procedures. They were instructed to write whatever they heard, regardless of whether it made sense. They could begin to type as soon as the speaker began talking, to offset issues with memory for the longer sentences. Speakers were split into two groups for this experiment, with half of the listeners hearing one group of four speakers (P1, P4, P7, P8), and the other half of listeners hearing the remaining four speakers P2, P3, P5, P6). This was done to ensure the experiment was of reasonable time length.

2.5.4 Rating Scale

Experiment Two consisted of listeners hearing the same sentence from the grandfather passage ('He slowly takes a short walk in the open air each day') from each of the eight speakers. They were asked to rate each speaker using a sliding scale as per Fletcher et al. (2017), with a range from 'difficult' at one end to 'easy' at the other. Listeners were asked to rate "how difficult is it to understand this speech" as a numeric output. This equated to most difficult represented by the number 0 and 1.0 representing speech that was very easy to understand.

2.5.5 Reaction Time Experiment

Experiment Three was the reaction time experiment. Two different sets of stimuli were used, each set consisting of four speakers with dysarthria and four healthy control speakers, all reading veracity statements. The control speakers in one stimuli set read identical statements to the speakers with dysarthria in the opposite stimuli set, so each listener heard every veracity statement from either a speaker with or without dysarthria. Listeners were instructed to respond the same way they did in the initial practice round, by pushing a button (n) if the statement was true and (m) for false. These buttons were assigned for their location on the keyboard, as fairly central for either hand. If they were unsure they were instructed to guess. Listeners heard 120 different statements, with their response time measured using custom-designed MATLAB software. Listeners who heard P1, P4, P7, and P8 in the Experiment One heard P2, P3, P5, and P6 in this experiment. This was to ensure each listener had measured each speaker in at least one form of assessment and to reduce perceptual learning effects (i.e. becoming more accurate or faster in processing a person's speech due to experience hearing the speaker in previous experiments).

2.5.6 Experiment Configuration

The three experiments were counterbalanced between participants, to avoid any effects of task order, which might affect the correlations between experiment types. The speakers were also split between the SIT and reaction time experiment (if the speaker's recordings were used in the SIT for one listener, the same speaker was not played in the reaction time experiment).

Finally, there was a fourth experiment where listeners heard three tones of different frequencies and pushed the spacebar once the tone was detected. This experiment was included as a baseline measure of reaction time to auditory stimuli and was completed last for all participants.

2.6. Data Treatment

The raw data from each listener was generated in an Excel output file, and these individual files were collated to create a master sheet for each experiment.

2.6.1 Reaction Time - Data Treatment

To calculate exact reaction time for each sentence, several measurements were required. The overall reaction time was recorded from the time the sentence began until the listener pushed a button to select their answer (true or false). However, each sentence was of a different recorded length, requiring individual measurements to be taken per sentence. For each sentence used, the point where a decision could be made to determine if it were true or false was identified and this point in the sentence was marked. For most sentences this point did not occur until the final word, as the pairing of the first and second word alone were ambiguous (e.g. Cars have ..., water is ...), but for a few sentences the middle word (a verb) could be used to determine if the sentence was false (e.g. ovens sew ..., clothes heat ...). Because spoken words undergo immediate processing by listeners—and can often be correctly identified prior to hearing the complete signal—the onset of the identified word was marked rather than the offset. Therefore, the time of possible sentence determination was identified at the onset of this word in each sentence, and the time from the start of the sentence until this point was subtracted from the overall reaction time to ensure that only the point of decision making was recorded.

There was also a tone based experiment (Experiment Four) where listeners pressed a button when a tone was played. This was done to account for how long it takes a listener to motorically respond to auditory stimuli. An average of this auditory reaction time for each listener was subtracted from their initial reaction time to each sentence. This was done to account for the variation in motoric movement, especially variation associated with age. It was assumed that any reaction to speech stimuli should be longer than this time, as during the

reaction time experiment the listener would also have to hear the final word of the of the sentence, process its meaning, and decide if the sentence was true or false, before they could motorically respond. The final, normalised measurement of sentence reaction time was defined as the time following the onset of the sentence's key "decisive" word, with the listener's average motor response time removed. This measurement was theorised to best reflect the reaction time that could be attributed to speech processing.

2.6.2 Reaction Time - Excluded Data

Any response time that was negative (e.g. quicker than the average motoric response to a pure tone) was removed, as this indicated that the listener had responded before hearing the sentence's key word and was thus assumed to be guessing. Any reaction time that was two standard deviations away from that listeners' average reaction time was also excluded, as it was assumed that these items reflected fatigue, guessing, or distraction during the task, as per Munro and Derwing (1995). This was between three and ten sentences per listener, from 120 sentences in total.

Each sentence spoken by a speaker with dysarthria was also spoken by a control speaker, to create paired sentences between groups. However, due to researcher error in the creation of the stimuli, from the 120 sentences used there were four that were not correctly matched. These were "queens are men" "swimming is exercise" "frogs take leaps" and "illustrators draw pictures." For analysis, these sentences were all excluded.

The accuracy of answers is not included as a factor in the overall analysis of reaction times. This is because the only way to account for accuracy would be to remove all sentences that did not receive over 75 percent accuracy across all speakers, to account for guessing on the part of the listener. Doing so would remove data that is still relevant for analysis of reaction time, which is the primary focus of this study. The sentences produced by speakers with low orthographic transcription scores were not expected to receive accurate listener

responses, as per the findings of Houben, van Doorn-Bierman & Dreshler (2013). They found an effect of reaction time regardless of how accurate participants were in completing the task, and it is expected in this study that longer response times will be found for severe speakers regardless of listener accuracy.

2.6.3 Orthographic Transcription – Data Treatment

Each target sentence and the number of target words correctly transcribed was recorded. For each sentence, the percentage correct was calculated based on the number of words correct out of the total number of target words. As per SIT protocols (Yorkston et al. 1996), any additions were ignored, and spelling mistakes and homonyms (words that sound the same but are spelt differently to the target word) were included as correct words. This resulted in a percent correct for each sentence, and these scores were averaged for each speaker to find the overall percent correct.

2.6.4 Rating Scale – Data Treatment

Ratings for each speaker by all listeners were tabulated into one spreadsheet. Possible scores ranged from 1.0 (representing easiest to understand) and 0.0 (most difficult speech to understand). For each rating, a z-score was calculated based on all scores provided by that listener. These z-scores were then averaged for each speaker to find their overall rating score. This was done as a measure to control for differences in how listeners used the scale.

2.7 Statistical Analyses

To address the research questions, several analyses were required. To investigate differences between the control and dysarthric group in the reaction time experiment, a comparison of means could not be done, as the data points were not normally distributed and there were many outliers. This is a common event with reaction time measures as the data is bound at one end (cannot be faster than zero) and can be theoretically any time beyond zero.

Despite removing data that was more than two standard deviations above or below the mean, there were still significant outliers within each group of dysarthria and control speakers that meant a parametric test could not be used. Instead, non-parametric analyses were completed.

A series of correlations were calculated to examine the relationships between the different tasks used to determine intelligibility. A multiple regression analysis was used to examine how hearing loss and age affected reaction time and orthographic intelligibility measures. These are further detailed in the results section below.

CHAPTER THREE:

RESULTS

The results are presented in four sections. The initial section describes the preliminary visualisation of the data, while the three main sections each address the individual research questions.

3.1 Initial Visualisation of Data

3.1.1 Reaction Time Data - Plotting and Tests of Normality

Younger Listeners

Figure 1 outlines the results of the reaction time data for younger listeners responding to both control and dysarthric speakers. As can be seen in Figure 1, the overall data is skewed.

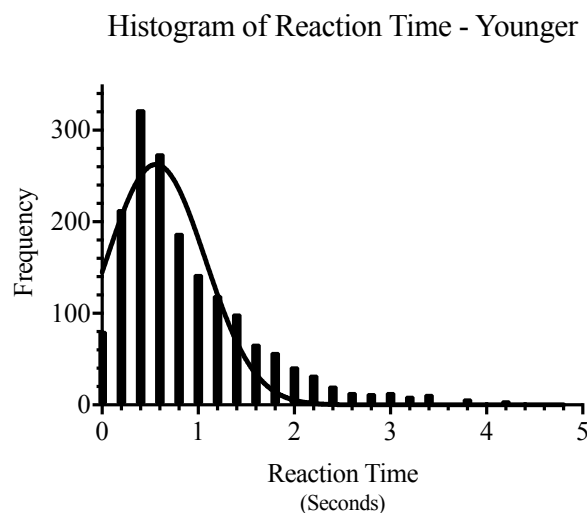


Figure 1. Bar chart frequency distribution of different reaction times to all speakers for younger listeners.

This is further demonstrated in Figure 2, comparing the younger listeners' reaction times to both the control speakers and speakers with dysarthria. As can be seen, a significant number of outlier values exist in listeners' reaction times to both control speech and those with dysarthria.

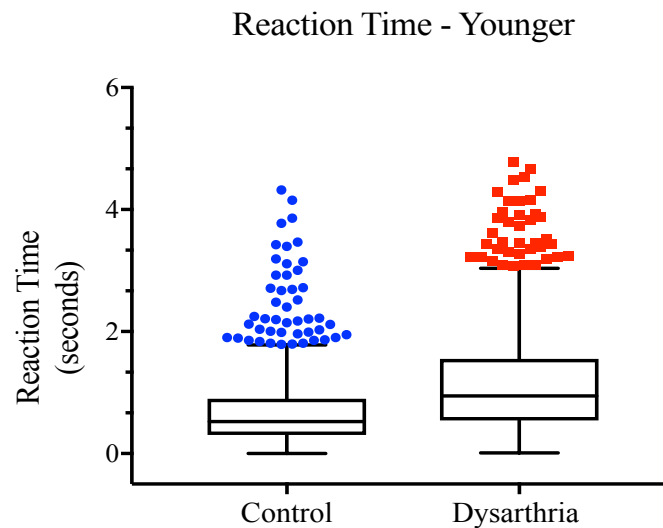


Figure 2. Mean (and interquartile ranges and extremes) reaction time for listeners to speakers with dysarthria and control speakers. Outliers identified with blue circles and red squares.

Overall, this preliminary analysis indicated a non-normal distribution of the data. Therefore, non-parametric tests were employed for all statistical comparisons of reaction time.

Older Listeners

The reaction time data from older listeners has a similar configuration to that of the younger listeners. Figure 3 highlights the overall skew of the data.

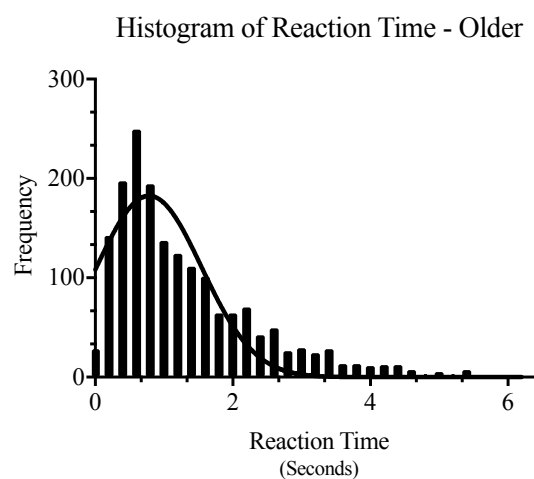


Figure 3. Bar chart frequency distribution of different reaction times to all speakers for older listeners.

As can be seen in Figure 4 there are again significant outliers in each group for control speakers (displayed in blue) and speakers with dysarthria (displayed in red).

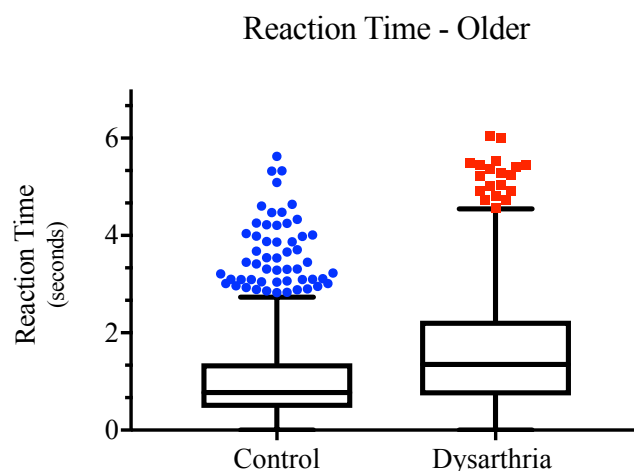


Figure 4. Box and whisker plot of mean (and interquartile ranges and extremes) reaction time for older listeners to speakers with dysarthria and control speakers. Outliers identified with blue (control) circles and red (dysarthria) squares.

As for the younger listeners' results, the need for non parametric analysis for reaction time is apparent from the distribution of data above.

3.1.2 Orthographic Transcription Data - Plotting and Tests of Normality

Younger Listeners

Figure 5 below shows the frequency distribution for transcription scores for younger listeners to all speakers with dysarthria.

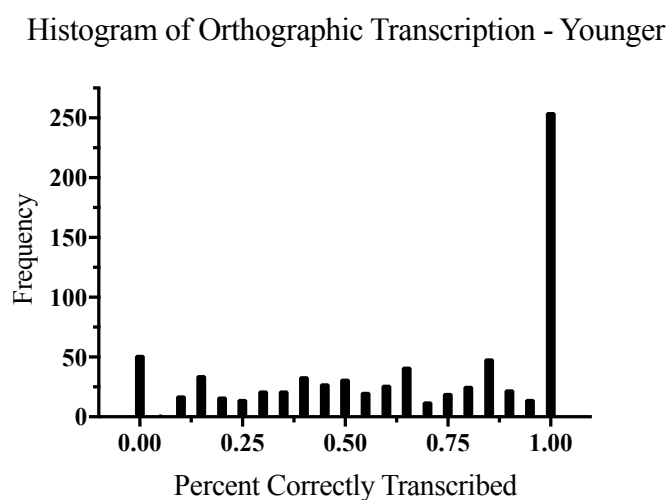


Figure 5. Bar chart frequency distribution of orthographic transcription scores to all speakers with dysarthria for younger listeners.

As with the reaction time data, a significant skew is present in this data, which means non-parametric testing is required for analysis.

Older Listeners

Figure 6 below shows the skew of orthographic transcription data from the older listeners.

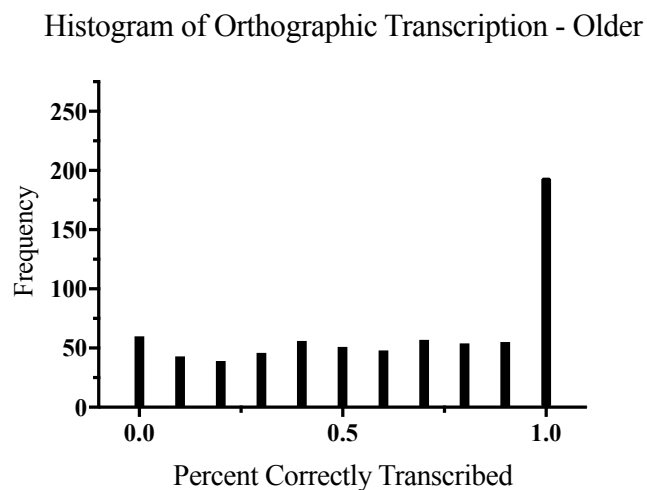


Figure 6. Bar chart frequency distribution of orthographic transcription to all speakers for older listeners.

These results are very similar to those from the younger listeners, and support the use of non parametric testing for this method of assessment.

3.1.3 Rating Scale Data - Plotting and Tests of Normality

Younger Listeners

Figure 7 reflects the distribution of rating score data from listeners in the younger age group.

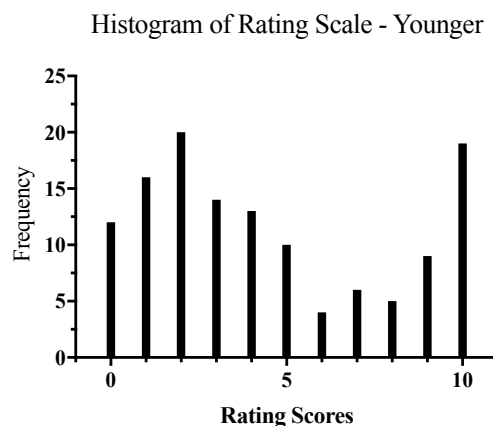


Figure 7. Bar chart frequency distribution of orthographic transcription scores to all speakers with dysarthria for younger listeners.

This figure illustrates the variability in the rating scale data. For consistency in comparisons between all three measures of dysarthria, non parametric testing is required for analysis.

Older Listeners

Figure 8 identifies the frequency distribution of scores in the rating scale component for older listeners.

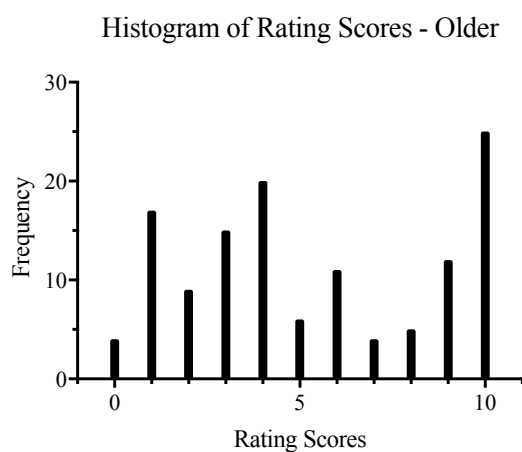


Figure 8. Bar chart frequency distribution of different reaction times to all speakers for older listeners.

As with the younger listeners, the rating scale data was varied and did not have sufficient data points to create a standard distribution.

3.2 Research Question One – Difference in Younger Listeners' Reaction Time to Dysarthric and Control Speakers

To investigate the first research question of whether reaction time can be used to distinguish between speakers with dysarthria and control speakers, younger speakers were focused on first. Figure 9 outlines the difference in reaction time for younger listeners hearing control speakers compared to speakers with dysarthria.

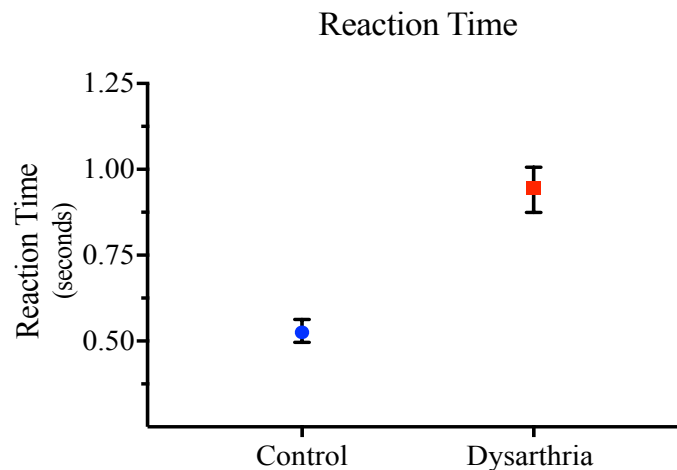


Figure 9. Median reaction time and 95 % confidence intervals displayed for control speakers and speakers with dysarthria, by younger listeners.

A Wilcoxon ranked sum test indicated that reaction time was significantly greater for speakers with dysarthria ($Mdn = 0.95$) than for control speakers ($Mdn = 0.53$) as $W = 236790$, $p < .001$, 95% CI = $[-0.42, -0.32]$, difference in location $-.37$, and $r = -.33$.

3.3 Research Question Two – Comparison of Reaction Time to Standard Measures

The second research question investigated similarities across different forms of dysarthria speech assessment. To compare the reliability of each assessment, the variability in listener scores was examined for each speaker with dysarthria within the reaction time experiment. The relationship was explored between individual speakers with dysarthria and their matched control speaker. The results are reported in Figure 10, with speakers with dysarthria paired with their control speaker adjacent to the right.

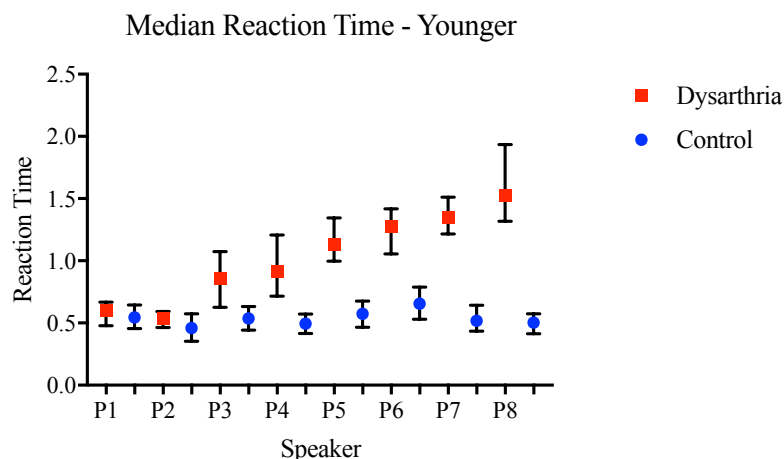


Figure 10. Median reaction time and 95% confidence intervals for individual speakers with dysarthria (red square) compared to their control speaker (blue circle) for younger listeners.

Control speakers had a relatively small range of median RTs (0.46 - 0.66) and have a smaller range in the 95% confidence intervals. Two speakers with mild dysarthria had a median value similar to controls (P1 and P2), while the remaining six pairs are significantly different.

3.3.1 Orthographic Transcription Analysis

Average percent intelligibility was calculated by finding the percent correct for each sentence and averaging the result for each speaker. This is displayed in Figure 11.

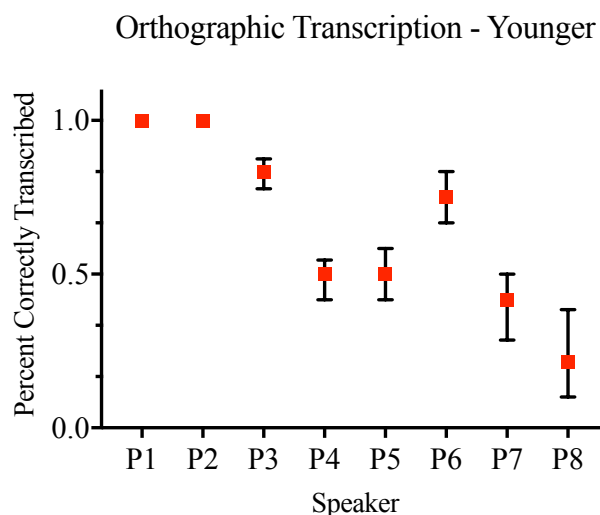


Figure 11. Median percent correct for orthographic transcription (1.0 = 100%, 0.5 = 50%) and 95% confidence intervals for individual speakers with dysarthria for younger listeners.

As displayed, the range in median values for transcription scores is from 22% correctly transcribed (for P8) to 100% correctly transcribed for both P1 and P2. The 95% confidence intervals are displayed for each speaker, further highlighting the range of scores for each speaker.

3.3.2 Rating Scale Analysis

The median rating score and confidence interval is displayed below in Figure 8. Distinct groupings have appeared, separating speakers into a more intelligible (P1, P2, P3) and a less intelligible cluster (P4, P5, P6, P7, P8).

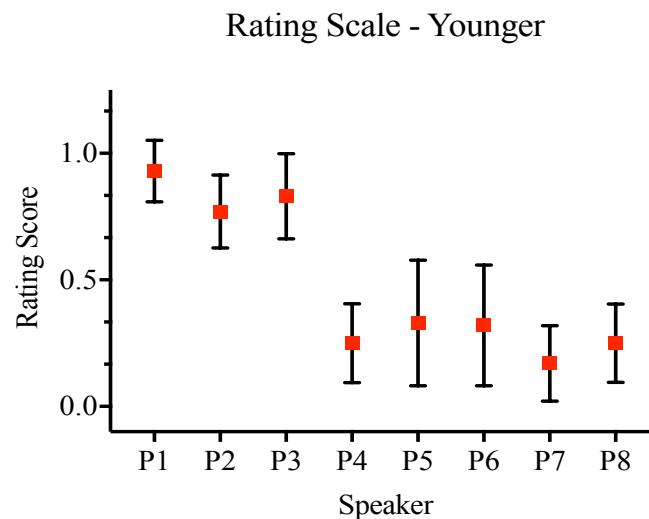


Figure 12. Median score and 95% confidence intervals for individual speakers with dysarthria by younger listeners. 0 = less intelligible, 1 = highly intelligible.

As with the orthographic transcription results, there is a range in average rating scores from 0.17 to 0.93, but the rating scores had more variability within each speaker, as evidenced by the 95% confidence intervals.

3.4 Comparisons Between Measures

After examining the variability in assessment scores within each speaker, the next aim was to compare the average score given to each speaker across different speech assessments. This allows further investigation into the convergent validity of these assessments.

3.4.1 Reaction Time vs Orthographic Transcription

When directly compared, figure 13 displays the significant correlation between reaction time results and percent correct in orthographic transcription for each speaker.

Reaction Time vs Orthographic Transcription - Younger

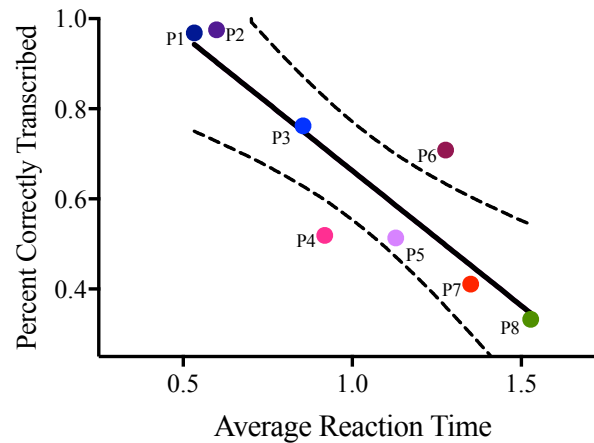


Figure 13. Percent correctly transcribed compared to average reaction time by speaker. Line of best-fit and dashed line representing 95% confidence interval for individual speakers.

This is a negative correlation, with $r = -.88$, $p = 0.004$, and 95% CI = $[-0.98, -0.46]$. This reflects that when more words are able to be transcribed, there is also a shorter reaction time.

3.4.2 Reaction Time vs Rating Score

Figure 14 represents the significant relationship found between reaction time and rating scale measures.

Reaction Time vs Rating Score - Younger

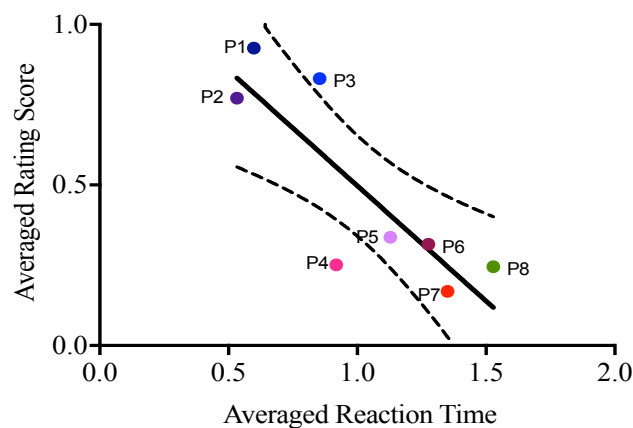


Figure 14. Average rating score compared to the average reaction time by speaker. Line of best-fit and dashed line representing 95% confidence interval marked.

Again, in this case, a negative correlation was the expected finding, with $r = -.84$, $p = .0047$, and 95% $CI = [-0.97, -0.33]$.

3.4.3 Rating Score vs. Orthographic Transcription

When comparing rating scores and percent of words correctly transcribed for each speaker, a significant correlation between these measures was identified. This is displayed below in Figure 15.

Orthographic Transcription vs Rating Scale - Younger

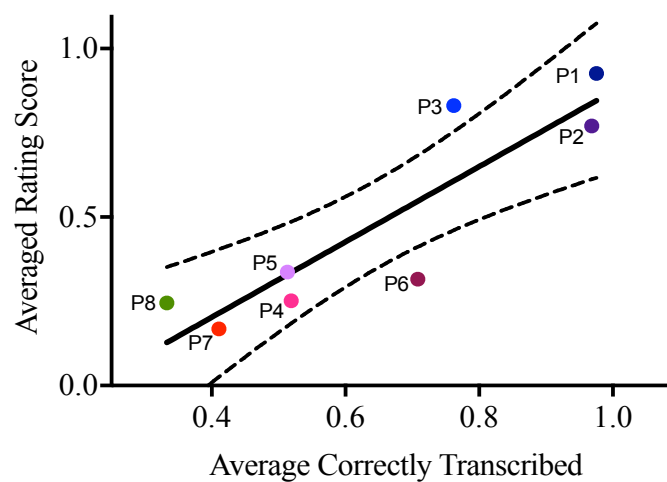


Figure 15. Percent correctly transcribed compared to rating score by speaker. Line of best-fit and dashed line representing 95% confidence interval.

A positive correlation was found with $r = .89$, $p = 0.002$ and 95% $CI = [0.49, 0.98]$, as a higher rating score corresponds to a higher percentage of correctly transcribed words.

3.5 Research Question Three - Effect of Age on Results

Our third research question was to establish if there is an effect of age and hearing loss on listener-based measurements of dysarthria. To address this, we first compared the median scores of listeners from both the younger and the older group when listening to speakers with dysarthria and control speakers.

3.5.1 Reaction Time Comparison

Figure 16 displays the reaction time results from both the younger and older groups when listening to speakers with dysarthria and the control speakers.

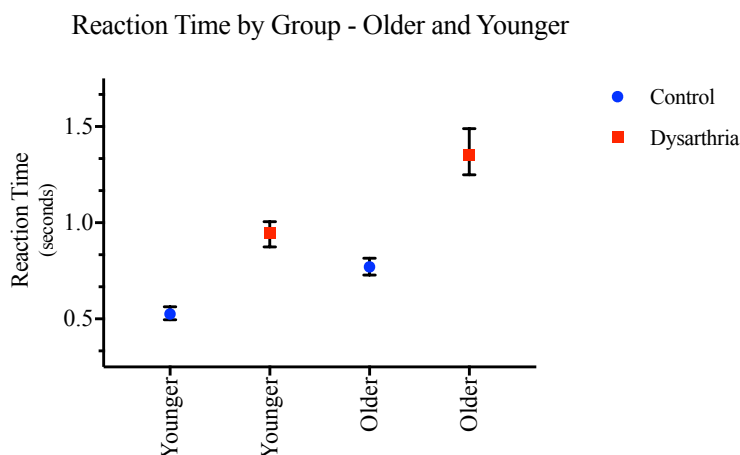


Figure 16. Median and 95% confidence interval displayed for reaction time by group to both control speakers (blue) and speakers with dysarthria (red).

As can be seen, younger listeners were typically faster to respond to both speakers with dysarthria and control speakers compared to older listeners, and this effect was more pronounced in the dysarthric group comparison.

Figure 17 further details the differences between the older and younger group, with reaction time median and 95% confidence intervals displayed for each speaker with dysarthria.

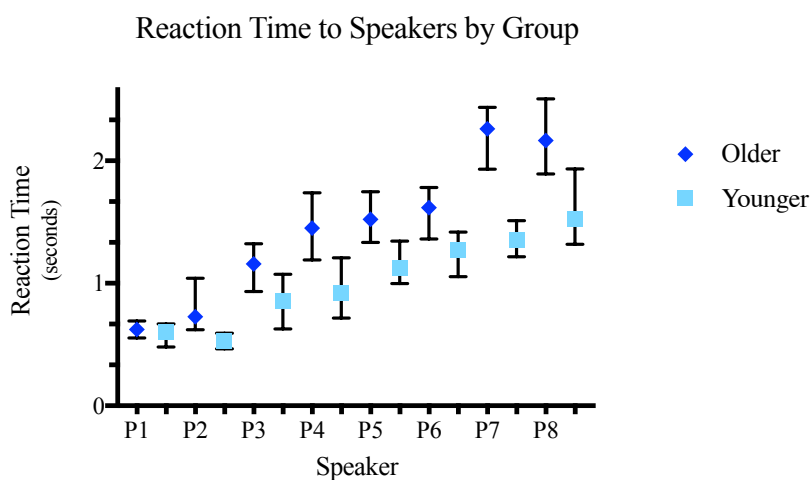


Figure 17. Median and 95% confidence interval displayed for reaction time to individual speakers with dysarthria between the older and younger groups.

This reflects a similar pattern of increasing reaction time for speakers who may be more severe in terms of speech degradation. For the milder speakers like P1, the two groups had very similar results, while a larger effect is seen for the more severe speakers like P7 and P8.

3.5.2 Effect of Age and Hearing Loss on Reaction Time

In measuring reaction time from older listeners, it is important to assess the effects of hearing loss and age on reaction time scores. To determine whether these factors had a significant effect on the data collected from this group, a backward-stepwise regression was calculated to predict reaction time based on age and hearing loss (by pure tone average) as well as the presence of dysarthria in the speech sample. This is displayed in figure 18.

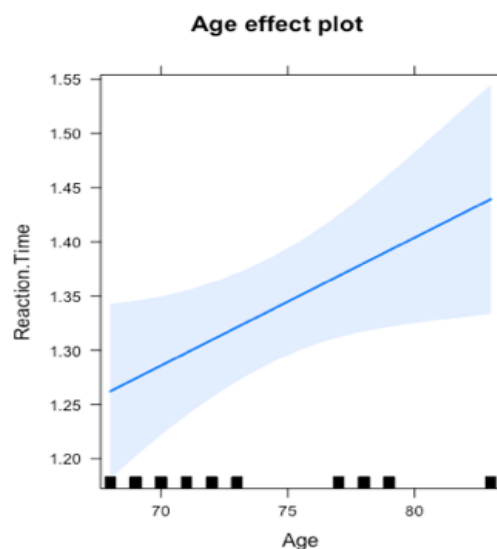


Figure 18. Linear regression of reaction time by age of listeners in the older group. 95% confidence intervals displayed in the blue shaded area.

The final model contained significant effects for age, pure tone average, presence of dysarthria and the interaction between pure tone average and presence of dysarthria ($F(4,1764) = 43.17, p < 0.001$) with an R^2 of 0.09.

Presence of dysarthria was a significant predictor of reaction time, with the reaction time of older listeners increasing by an estimated average of .93 seconds when listening to someone with dysarthria compared to a control speaker. As demonstrated in Figure 18, as age

increases, reaction time also significantly increases with this model. This is apparent in the comparison between the older and younger groups, but it is also true within the older group, where age ranged from 69 to 83. The oldest participants had the slowest reaction times when hearing loss was controlled. Age accounted for a .01 of a second delay in reaction time as participants aged per year.

Another finding from this model was the effect that hearing loss had on listeners' reaction time to control speakers, as demonstrated by Figure 19.

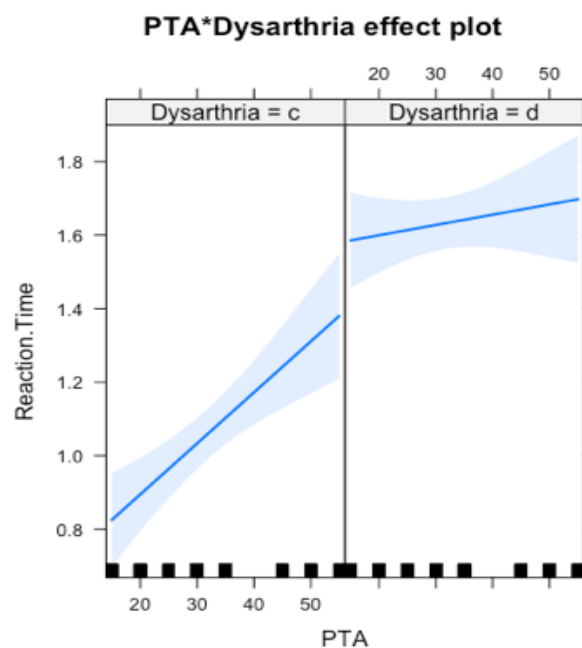


Figure 19. Regression of effect of pure tone average on reaction time. Control speakers on the left (dysarthria = c) and speakers with dysarthria on the right (dysarthria = d).

When listeners' pure tone average increased, so too did the amount of time needed to react to a speaker. There was a significant interaction between pure tone average and reaction time, with a one dB increase in hearing loss responsible for a 0.01 second increase in reaction time. This effect was much stronger for the control group than for the speakers with dysarthria.

3.5.3 Effect of Age and Hearing Loss on Orthographic Transcription

In addition to having slower reaction times, it was also hypothesized that older listeners would have lower accuracy in transcribing sentences from speakers with dysarthria. The older group were reviewed using a backward-stepwise regression to predict percent correct in the sentence intelligibility test by hearing loss and age, and the results are displayed in Figure 20.

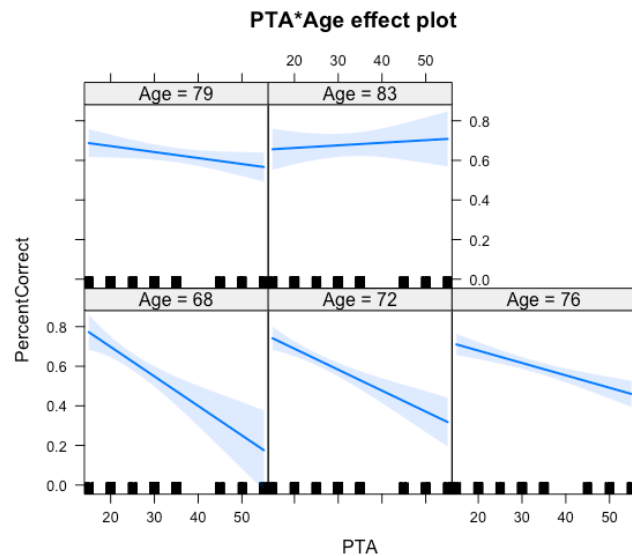


Figure 20. Regression of effect of pure tone average on percent correct controlling for age. age 68-71 n =seven, age 72 -75 n= two, 76 -78 n= four, 79 – 82 n = one and 83 n = two.

A significant regression equation was found ($F(3, 700) = 9.011, p < .001$), with an R^2 of 0.03. Based on these results, it appears that as pure tone average increases, percent correct decreases, with an 8% drop in accuracy per dB step in pure tone hearing threshold. Similarly, as age increases, percent correct also decreases, however age carries a smaller effect, with 2% poorer transcription scores for each year of increased age.

Differences exist between each age bracket described in Figure 19, in terms of how hearing loss affects percent correct in orthographic transcription. For participants aged 68-71, it appears that as pure tone average increases, there is a larger decrease in percent correct compared to other age groups, however, the 95% confidence interval (represented as the blue shaded area) becomes significantly wider for greater pure tone averages. This is extrapolated data from groups with small sample sizes however, and must be interpreted with caution.

CHAPTER FOUR:

DISCUSSION

The purpose of this research study was to see if reaction time is a viable method of assessing dysarthria, that may address the limitations current measures contain. In addition, this study examined the extent to which these assessments were affected by listeners' age and degree of hearing loss. Listeners heard veracity statements from people with a range of dysarthria severities and control speakers and responded by pushing a button to indicate their answer, which was timed. It was hypothesized that there would be a significant difference in reaction time for control speakers and speakers with dysarthria. Two groups were tested, one containing listeners aged between 18-30, and an older group aged 65 and over.

4.1 Overview of Results

Three main research questions were addressed in this study, and the findings of each question are detailed below.

4.1.1 Difference in Reaction Time to Speakers with Dysarthria and Control Speakers.

As can be seen in the results section above, reaction times were significantly different between control speakers and speakers with dysarthria for younger listeners. This confirms the first hypothesis, that there would be a significant difference between groups, with it taking significantly longer for adults to respond to speech from someone with dysarthria. This is consistent with work from Munro & Derwing (1995) and Evitts & Searle (2016), who both found significant differences in their comparisons of speakers with accents or dysphonia respectively, and control speakers.

When this was further analysed, looking at individual pairings of control speakers and speakers with dysarthria, this difference in reaction time was much greater for speakers with

severe dysarthria compared to mild speakers. Reaction time appeared to be fairly consistent between all control speakers at an average of 0.53 seconds, and the mildest speakers with dysarthria who were slightly longer than this. Given how small these differences were, and the general variability in listener responses, the differences were not statistically significant. To determine whether this effect was consistent and significant, many more listeners would be required. This could clarify if there is a floor effect with reaction time.

As Figure 6 illustrates, reaction time does become greater as a speaker is more affected by dysarthria, confirming the second hypothesis, that severity of dysarthria would lead to an increase in reaction time. This indicates that it is possible that reaction time could also be used to indicate the degree of severity in speakers if it were further investigated and standardized.

Increased reaction time to less intelligible speech was consistent with findings from Wilson and Spaulding (2010) and their three groups, one native English speaking (control group), and two groups of Korean accented English speakers, one more intelligible than the other. Reaction time latencies increased as fluency in English decreased, illustrating a change in reaction time related to degree of processing ease.

This finding is expected, as it aligns with the knowledge that listeners' reaction times have consistently been higher when exposed to more challenging listening conditions, (Gough 1965; Gatehouse & Gordon, 1990; Munro & Derwing, 1995; Wilson & Spaulding, 2010; Houben, van Doorn-Bierman & Dreshler, 2013). Prolonged reaction time has been used numerous times as evidence of increased cognitive difficulty and processing time. It reflects the added cognitive resources needed to attend to imprecise stimuli, and possibly the added requirement from short term and working memory to review what was heard before acting on the message (Munro & Derwing, 1995).

4.1.2 Reaction Time Compared to Standard Measures in Younger Listeners

In terms of how well this measure compares to previously accepted forms of dysarthria measurement, reaction time had a strong correlation to orthographic transcription accuracy in younger listeners, accounting for 88% of the variance in the average scores given to each speaker. Reaction times were also closely associated with rating scores, accounting for 84% of the inter-speaker variance in these measurements. The rating scale and transcription results were similarly strongly correlated, at 89%. The fact that all three measures were closely associated provides evidence for the reaction time experiment to be considered as a potential tool for further investigation as a measure of dysarthria severity. This finding is also consistent with previous comparative literature of both rating scales and orthographic transcription, and aligns with the research by Munro and Derwing (1995) and Wilson and Spaulding (2010) who used components of orthographic transcription with their reaction time assessments.

4.1.3 Effect of Age on Findings

A. Reaction Time

When the reaction time experiment was repeated with an older listener group, similar results were found. There was still a significant difference between control speakers and speakers with dysarthria as a group, however the overall reaction time latencies to both groups were longer for the older listeners. There remained a noticeable increase in reaction time between speakers with increased severity of dysarthria, with a lack of distinction between the milder speakers and control speakers. It is promising that this similarity remains, as it indicates that longer reaction time to more severe speakers occurs regardless of age, and may reflect additional difficulty with processing severely affected speech.

When hearing loss and age were examined further within the older adults, some interesting findings emerged. Within the older group, age was a significant factor in reaction

time, with an increase in listener age significantly related to an increased reaction time. While reaction time to speakers with dysarthria remained fairly consistent despite an increase in hearing loss, a higher pure tone average was significantly associated with a longer reaction time to control speakers.

The disproportionate effect of hearing loss on listeners' reaction time to healthy speakers was an unanticipated and interesting finding. It is likely that some of the acoustic speech features that are used to rapidly match stimuli to their lexical representations are not perceived by listeners with hearing loss. This may disproportionately affect their perception of healthy speakers, who accurately produce many distinctive acoustic features that aid in this matching process. In contrast, when acoustic features are not accurately produced in dysarthria, reaction time might not rely quite so heavily on minor differences in hearing acuity. Hence, for someone with a higher pure tone average, speech from a control speaker may only be providing a comparable amount of information as speech from a person with dysarthria to decode the message.

Another theory as to why there is no steady increase in reaction time with hearing loss could be that there is a reaction time ceiling effect. After approximately 1.5 seconds (to consider what they heard), people who are unsure of what they heard will have to guess, regardless of how well they heard the signal. The reason for this uncertainty could be due to the lack of clarity from a speaker with more severe dysarthria, the severity of the listeners hearing loss, or a combination of the two. This would mean that reaction time would not continue to increase with increased hearing thresholds, as there would be a point (possibly around 1.5 seconds) where listeners would have selected their answer, regardless of their confidence in choice.

These findings confirm the hypothesis that age is be a significant factor in reaction time. Hearing loss is also a significant factor in reaction time, but larger effects were demonstrated with the control speakers than the speakers with dysarthria.

B. Orthographic Transcription

When age was considered as a factor in transcription scores, the hypothesis was again confirmed that age would have a significant effect on scores. Not only were the older group less accurate when compared to the younger listeners, but there was a continuing decline accuracy as age increased. Of interest, hearing loss accounted for more variance in transcription accuracy than age alone. It appears that as both pure tone average and age increases, percent correct decreases, but it was the increase in hearing loss that contributed to the transcription ability the most. These findings are consistent with research into the effects on age on communication and processing. Working memory and related cognitive processes can be affected by age (Verhaeghen & Salthouse, 1997). As stated above, this is likely due to the combined effect of having a speech signal that is degraded to a varying extent by the speaker, and the listener not being able to receive a clear message due to the loss of hearing at important speech frequencies. This dual effect on the message clarity is reflected in the ability of these listeners with the transcription task. The fact that the older listeners transcribed fewer correct sentences than the younger listeners is not surprising, but the additional information about the effects of hearing loss on transcription scores combined with age effects does provide scope for further investigation in this field.

4.2 Study Limitations

While every effort was made to anticipate and prepare for the possible difficulties that would occur during this study, several limitations exist that have either affected the findings of the study or, if they had been better addressed at the time, could have strengthened the

conclusions drawn from this data. These have been identified below and ideas for how they could be addressed in future studies are included.

4.2.1 Veracity Statements – Stimuli and Data Collection

One limitation is that the speakers with dysarthria were recorded separately from the control speakers. A different researcher made the control recordings and they were conducted in a different space to where the original recordings were made. As discussed by Schiavetti (1992), variation in the transmission system/stimuli can result in uncontrolled variation in the results of a study. In future studies, accuracy could be increased by having more consistency in the recordings of the stimuli.

Another aspect of this study that could have been improved was the choice of veracity statements. For this test to be universally applicable to all listeners of a variety of ages and vocabulary levels, there should be further testing done to ensure these sentences do not offer more variation than is needed based on listener knowledge. A sentence that is more complex in vocabulary may require a longer amount of time to process, affecting the results of a reaction time test. It would have been interesting to have used the same statements as previous researchers such as Munro and Derwing, (1995) and Wilson and Spaulding (2010) who used the same veracity statements. This could have allowed for more consistency in comparisons with the current literature.

4.2.2 Scoring of Veracity Statements

Because a veracity statement inherently only has two possible answers, true or false, this leaves a 50% chance that a participant who is unsure of the correct answer will still guess the correct answer.

For the present study, reaction time was the dependent variable of interest. Accuracy also could have been a valid dependent variable if a way of determining accuracy to rule out

guesswork had been determined, but this was not done before testing began. This was not required to adequately answer the primary research question of the paper, but including this could have added more information for this study to report.

Something that could have been done to improve the interpretation of results would have been to replicate what Munro and Derwing (1995) did in their study comparing reaction time in response to native speakers of English and Mandarin-accented English speakers.

Listeners would hear all speakers saying a sentence in English, and select between either 'true' or 'false'. They would then write down what they heard and move onto the next sentence. This allowed the researchers to ensure that responses were accurate, as they could compare their veracity response with what they had transcribed, to ensure it matched and could eliminate sentences that were not correctly transcribed.

Adding this feature to the present study would have provided a different way to analyse the results. Being able to definitively determine accuracy would have been beneficial, as further analysis looking at accuracy compared to reaction time could be conducted.

4.2.3 Limited Data Collection

There is additional data that could have been collected from listeners to better interpret their responses. Factors like memory and vocabulary have been proven in previous research to affect outcomes in tasks involving listening to speakers with dysarthria.

McAuliffe, Gibson, Kerr, Anderson, & LaShell, (2013) found that individuals with a larger receptive vocabulary were better able to interpret speech cues from speakers with dysarthria.

Another key factor to better intelligibility results by listeners is working memory ability. Baddeley and Hitch (1974) stated that working memory is the term used to describe a person's ability to store and use linguistic information. This is done by focusing on the

important elements of the signal and ignoring extraneous information. People with dysarthria produce speech that contains many elements of distraction or difference compared to clear speech, so speech may need to be stored for longer in the listeners working memory to focus on the specific words that have been said.

Both the reaction time task and orthographic transcription task rely on recognition of words, and a more robust vocabulary and memory would logically make these tasks easier.

Though it would increase the amount of time required for each participant to be tested, it would be worthwhile for a future study to repeat the experiments and include a robust measure of receptive vocabulary and memory.

4.2.4 Summary of Limitations

Veracity stimuli appear to be a good tool for examining processing speed of speech. It would be beneficial if there were a standardized set of stimuli, where sentences are universally used, that do not require an extensive vocabulary to accurately participate. It could be beneficial to combine components of the orthographic transcription task with the reaction time measure, to create a test that more accurate and efficient in determining the severity of dysarthria.

Ultimately, developing this assessment as a tool for measuring dysarthria severity will require more data. Repeating this experiment to check the reliability of results would be required, and with the procedural errors removed and limitations reduced the variation between listener responses is likely to be reduced. Further listener factors could be assessed such as working memory and vocabulary, to include additional predictors of performance and understand what will affect the listener's ability to perform this task.

4.3. Summary of Findings

Overall a significant difference between reaction time to speakers with dysarthria compared to control speakers was found. This effect was strongest between speakers more

severely affected by dysarthria, with no significant difference identified between the pairs of more mild speakers and their controls. As a measure, reaction time correlated well with previously established measures of dysarthria severity like orthographic transcription and rating scales. When the results from the younger group were compared to those of the older adults tested, a similar pattern of results emerged. Hearing loss and age as factors contributed to performance in a small, but significant, way on both transcription scores and reaction time latencies when assessed using a regression model. This study is the preliminary step in the process of changing dysarthria assessment, and as demonstrated above, there is convincing evidence that this must be further investigated.

4.4 Future Research Directions

There is compelling evidence to support further investigation into reaction time (or a combined test) as a replacement of orthographic transcription as the gold standard for speech assessment of dysarthria. As an assessment, this has the potential to be much quicker and less demanding than using orthographic transcription alone and could be a more consistent and repeatable measure than a rating scale. The reaction time test is also the most transferable to a real-world context. It involves listening to a sentence, comprehending that sentence then responding appropriately. This simulates daily interactions more closely than repeating what someone has said.

Developing this form of testing to become a standardised measure would require a study similar to what has been conducted in this research, using a more robust design. Several of the limitations discussed above can be addressed with changes to the methodology of the experiment, and by having more listeners and speakers, to give more power to this study. The preliminary findings of this study do indicate that control speakers generate a similar reaction time across all speakers, and if normative data could be collected, comparisons can be drawn between speakers and the standard time recorded for controls.

Creating a combined test of a reaction time to a veracity statement followed by transcription of the statement clarifies several current limitations of accuracy measurement, and also provides a closed and open set framework within the one test. Having a closed set with just the two options, true or false, helps to negate response bias, as listeners must choose an answer regardless of their confidence, while the transcription, as an open set response, provides more insight into accuracy and information about the type of errors made.

Investigation into the effect of hearing loss is an area that is often neglected in studies of dysarthria ratings. Based on the prevalence of both dysarthria and hearing loss within the elderly population, it is a necessary area of future study. Adding data about effects of hearing loss would provide insight into the additional difficulty faced by communication partners when there is both hearing loss and dysarthria affecting communication. As the preliminary results from this study suggest that hearing loss associated with aging produces a significant difference in reaction times, further research into communication between partners with both issues present is required. A study with a larger number of older listeners separated into two groups, normal hearing and those with hearing loss could provide more information about the role hearing loss has in speech understanding of speakers with dysarthria.

4.5 Clinical Implications

For reaction time to be used clinically, clinicians would need to be able to administer this test and score it quickly, and it is likely that they would be the one responding to the veracity statements. If a person was very familiar with the veracity statements, they would be able to accurately respond before the sentences has been fully spoken by the speaker with dysarthria.

A way to overcome this issue would be for each true veracity statement to have a nearly identical false statement, that differs in the final word. That way the listener won't know if it is the true or false version of that sentence until the onset of the final word (e.g.

trees have leaves, trees have shoes). That way a small list of sentences randomly read by the speaker could still be responded to by a clinician familiar with the test sentences. A reaction time measure can be easily imagined as part of the test battery for a clinician when completing communication assessments.

Ultimately, there is compelling evidence to support reaction time measures as a useful clinical tool in assessment of dysarthria. This form of assessment is quick to administer, and is directly representative of daily communicative ability of the speaker. With more research focus on stimuli and procedures, this could serve as a model for not only assessment of dysarthria, but for language feature analysis in many other domains.

References

- Anand, S., & Stepp, C. E. (2015). Listener Perception of Monopitch, Naturalness, and Intelligibility for Speakers with Parkinson's Disease. *Journal of Speech Language and Hearing Research*, 58(4), 1134-1144.
- Bashore, T. R., Osman, A., & Heffley, E. F. (1989). Mental Slowing in Elderly Persons - a Cognitive Psychophysiological Analysis. *Psychology and Aging*, 4(2), 235-244.
- Bassey, E. J., & Harries, U. J. (1993). Normal values for handgrip strength in 920 men and women aged over 65 years, and longitudinal changes over 4 years in 620 survivors. *Clinical Science*, 84, 331-337.
- Birren, J. E., Woods, A. M., & Williams, M. V. (1979). Speed of behavior as an indicator of age changes and the integrity of the nervous system. In F. Hofmeister & C. Mueller (Eds.), *Brain functions in old age* (pp. 10-44). Berlin: Springer-Verlag.
- Borrie, S. A., McAuliffe, M. J., Liss, J. M., Kirk, C., O'Beirne, G. A., & Anderson, T. (2012). Familiarisation conditions and the mechanisms that underlie improved recognition of dysarthric speech. *Language and Cognitive Processes*, 27(7-8), 1039-1055.
- Botwinick, J., & Thompson, L. W. (1966). Components of reaction time in relation to age and sex. *The Journal of genetic psychology*, 108(2), 175-183.
- Bradburn, N. M., & Miles, C. (1979). Vague quantifiers. *Public Opinion Quarterly*, 43(1), 92-101.
- Bunton, K., Kent, R. D., Duffy, J. R., Rosenbek, J. C., & Kent, J. F. (2007). Listener agreement for auditory-perceptual ratings of dysarthria. *Journal of Speech, Language, and Hearing Research*, 50(6), 1481-1495.
- Cerella, J., Poon, L. W., & Williams, D. M. (1980). Age and the complexity hypothesis. In L. W. Poon (Ed.), *Aging in the 1980s* (pp. 332-340). Washington, DC: American Psychological Association.
- Committee on Hearing, Bioacoustics, and Biomechanics. (1988). Speech understanding and aging. *Journal of the Acoustical Society of America*, 83, 859-895.
- Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech, Language, and Hearing Research*, 12(2), 246-269.

- Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders*. Philadelphia: Saunders.
- Divenyi, P. L., Stark, P. B., & Haupt, K. M. (2005). Decline of speech understanding and auditory thresholds in the elderly. *Journal of the Acoustical Society of America*, 118, 1089–1100.
- Divenyi, P. L., Stark, P. B., & Haupt, K. M. (2005). Decline of speech understanding and auditory thresholds in the elderly. *Journal of the Acoustical Society of America*, 118, 1089–1100.
- Dos Santos Barreto, S., & Ortiz, K. Z. (2015). Protocol for the Evaluation of Speech Intelligibility in Dysarthrias: Evidence of Reliability and Validity. *Folia Phoniatrica et Logopaedica*, 67(4), 212-218.
- Duffy, J. R. (2006). History, Current Practice and future trends and goals. In Gary Weismer (Eds.) *Motor Speech Disorders: Essays for Ray Kent* (pp. 7-56). San Diego: Plural Pub.
- Evitts, P. M., & Searl, J. (2006). Reaction times of normal listeners to laryngeal, alaryngeal, and synthetic speech. *Journal of Speech, Language, and Hearing Research*, 49(6), 1380-1390.
- Evitts, P. M., Starmer, H., Teets, K., Montgomery, C., Calhoun, L., Schulze, A., . . . Adams, L. (2016). The Impact of Dysphonic Voices on Healthy Listeners: Listener Reaction Times, Speech Intelligibility, and Listener Comprehension. *American Journal of Speech-Language Pathology*, 25(4), 561-575.
- Finkel, D., Reynolds, C. A., McArdle, J. J., Hamagami, F., & Pedersen, N. L. (2009). Genetic variance in processing speed drives variation in aging of spatial and memory abilities. *Developmental psychology*, 45(3), 820.
- Fletcher, A. R., McAuliffe, M. J., Lansford, K. L., & Liss, J. M. (2017). Assessing Vowel Centralization in Dysarthria: A Comparison of Methods. *Journal of Speech, Language, and Hearing Research*, 60(2), 341-354.
- Frontera, W. R., Hughes, V. A., Fielding, R. A., Fiatarone, M. A., Evans, W. J., & Roubenoff, R. (2000). Aging of skeletal muscle: A 12-yr longitudinal study. *Journal of Applied Physiology*, 88, 1321–1326.
- Gates, G. A., & Mills, J. H. (2005). Presbycusis. *The Lancet*, 366(9491), 1111-1120.
- Gough, P. B. (1965). Grammatical transformations and speed of understanding. *Journal of verbal learning and verbal behavior*, 4(2), 107-111.
- Hustad, K. C. (2006). A closer look at transcription intelligibility for speakers with dysarthria: Evaluation of scoring paradigms and linguistic errors made by listeners. *American Journal of Speech-Language Pathology*, 15(3), 268-277.

- Hustad, K. C. (2008). The relationship between listener comprehension and intelligibility scores for speakers with dysarthria. *Journal of Speech Language and Hearing Research*, 51(3), 562-573.
- Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of experimental psychology*, 45(3), 188.
- Kallman, D. A., Plato, C. C., & Tobin, J. D. (1990). The role of muscle loss in the age-related decline of grip strength: Crosssectional and longitudinal perspectives. *Journal of gerontology*, 45(3), M82-M88.
- Kangas, K. A., & Allen, G. D. (1990). Intelligibility of synthetic speech for normal-hearing and hearing-impaired listeners. *Journal of Speech and Hearing Disorders*, 55(4), 751-755.
- Kim, Y., Kent, R. D., & Weismer, G. (2011). An acoustic study of the relationships among neurologic disease, dysarthria type, and severity of dysarthria. *Journal of Speech, Language and Hearing Research*, 54, 417.
- Kreul, E. J. (1972). Neuromuscular control examination (NMC) for Parkinsonism: Vowel prolongations and diadochokinetic and reading rates. *Journal of Speech, Language, and Hearing Research*, 15(1), 72-83.
- Lee, Y., Sung, J. E., & Sim, H. (2014). Effects of listeners' working memory and noise on speech intelligibility in dysarthria. *Clinical Linguistics & Phonetics*, 28(10), 785-795.
- Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, 6, 172-187.
- Liss, J. M., Spitzer, S., Caviness, J. N., Adler, C., & Edwards, B. (1998). Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech. *Journal of the Acoustical Society of America*, 104, 2457-2466.
- Liss, J. M., Spitzer, S. M., Caviness, J. N., Adler, C., & Edwards, B. W. (2000). Lexical boundary error analysis in hypokinetic and ataxic dysarthria. *Journal of the Acoustical Society of America*, 107, 3415-3424.
- Liss, J. M. (2007). The role of speech perception in motor speech disorders. In G. Weismer (Ed.), *Motor speech disorders: Essays for Ray Kent* (pp. 195-231). San Diego, CA: Plural.
- Luciano, M., Gow, A. J., Harris, S. E., Hayward, C., Allerhand, M., Starr, J. M., ... & Deary, I. J. (2009). Cognitive ability at age 11 and 70 years, information processing speed, and APOE variation: the Lothian Birth Cohort 1936 study. *Psychology and aging*, 24(1), 129.

- Lyons, J., Elliott, D., Swanson, L. R., & Chua, R. (1996). The use of vision in manual aiming by young and older adults. *Journal of Aging and Physical Activity*, 4, 165–178.
- McAuliffe, M. J., Borrie, S. A., Good, P. V., & Hughes, L. (2010). Consideration of the listener in the assessment and treatment of dysarthria.
- McAuliffe, M. J., Gibson, E. M. R., Kerr, S. E., Anderson, T., & LaShell, P. J. (2013). Vocabulary influences older and younger listeners' processing of dysarthric speech. *The Journal of the Acoustical Society of America*, 134(2), 1358-1368.
- Moscicki, E. K., Elkins, E. F., Baur, H. M., & McNamara, P. M. (1985). Hearing loss in the elderly: an epidemiologic study of the Framingham Heart Study Cohort. *Ear and Hearing*, 6(4), 184-190.
- Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and speech*, 38(3), 289-306.
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., ... & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695-699.
- Patel, R., Usher, N., Kember, H., Russell, S., & Laures-Gore, J. (2014). The influence of speaker and listener variables on intelligibility of dysarthric speech. *Journal of Communication Disorders*, 51, 13-18.
- Reynolds, M. E., & Fucci, D. (1998). Synthetic Speech Comprehension A Comparison of Children With Normal and Impaired Language Skills. *Journal of Speech, Language, and Hearing Research*, 41(2), 458-466.
- Salthouse, T. A. (1985). Speed of behavior and its implications for cognition. In J. W. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 400-426). New York: Van Nostrand Reinhold.
- Schiavetti, N., & Kent, R. D. (1992). Scaling procedures for the measurement of speech intelligibility. *Intelligibility in speech disorders*, 11-34.
- Street Jr, R. L., Brady, R. M., & Putman, W. B. (1983). The influence of speech rate stereotypes and rate similarity on listeners' evaluations of speakers. *Journal of Language and Social Psychology*, 2(1), 37-56.

- Sussman, J. E., & Tjaden, K. (2012). Perceptual measures of speech from individuals with Parkinson's disease and multiple sclerosis: Intelligibility and beyond. *Journal of Speech, Language, and Hearing Research*, 55(4), 1208-1219.
- Tikofsky, R. S., & Tikofsky, R. P. (1964). Intelligibility measures of dysarthric speech. *Journal of Speech, Language, and Hearing Research*, 7(4), 325-333.
- Tikofsky, R. S. (1970). A revised list for the estimation of dysarthric single word intelligibility. *Journal of Speech, Language, and Hearing Research*, 13(1), 59-64.
- Tjaden, K., & Wilding, G. E. (2004). Rate and loudness manipulations in dysarthria: Acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research*, 47(4), 766-783.
- Turner, G. S., Tjaden, K., & Weismer, G. (1995). The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*, 38, 1001-1013.
- Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age-cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin*, 122, 231-249.
- Wechsler, D. (1981). *WAIS-R manual: Wechsler adult intelligence scale-revised*. Psychological Corporation.
- Weinstein, B. E. "Hearing Loss in the Elderly: A New Look at an Old Problem." *Handbook of Clinical Audiology*. Ed. Jack Katz. 7th ed. Baltimore: Williams and Wilkins, 2014. 631-46.
- Wilson, E. O. B., & Spaulding, T. J. (2010). Effects of noise and speech intelligibility on listener comprehension and processing time of Korean-accented English. *Journal of Speech, Language, and Hearing Research*, 53(6), 1543-1554.
- Weismer, G., Kent, R. D., Hodge, M., & Martin, R. (1988). The acoustic signature for intelligibility test words. *The Journal of the Acoustical Society of America*, 84(4), 1281-1291.
- Weismer, G., Jeng, J.-Y., Laures, J. S., Kent, R. D., & Kent, J. F. (2001). Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. *Folia Phoniatrica et Logopaedica*, 53, 1-18.

- Weismer, G., & Laures, J. S. (2002). Direct Magnitude Estimates of Speech Intelligibility in Dysarthria Effects of a Chosen Standard. *Journal of Speech, Language, and Hearing Research*, 45(3), 421-433.
- Weiss, B.D., & Coyne, C. (1997). Communicating with patients who cannot read. *The New England Journal of Medicine*, 337, 272-274.
- Yorkston, K. M., Beukelman, D. R., & Tice, R. (1996). Sentence intelligibility test. *Lincoln, NE: Tice Technology Services*.
- Yorkston, K. M., & Beukelman, D. R. (1981). Communication efficiency of dysarthric speakers as measured by sentence intelligibility and speaking rate. *Journal of Speech and Hearing Disorders*, 46(3), 296-301.

Appendices

Appendix A

Consent from University of Canterbury Human Ethics Committee



HUMAN ETHICS COMMITTEE

Secretary, Rebecca Robinson
Telephone: +64 03 369 4588, Extn 94588
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2017/37/LR

3 July 2017

Rebecca Risi
Communication Disorders
UNIVERSITY OF CANTERBURY

Dear Rebecca

Thank you for submitting your low risk application to the Human Ethics Committee for the research proposal titled "Examining Measurements of Speech Disorder".

I am pleased to advise that this application has been reviewed and approved.

Please note that this approval is subject to the incorporation of the amendments you have provided in your email of 21st June 2017.

With best wishes for your project.

Yours sincerely

R. Robinson
pp.

Associate Professor Jane Maidment
Chair, Human Ethics Committee



HUMAN ETHICS COMMITTEE

Secretary, Rebecca Robinson
Telephone: +64 03 369 4588, Extn 94588
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2017/37/LR Amendment 1

18 July 2017

Rebecca Risi
Communication Disorders
UNIVERSITY OF CANTERBURY

Dear Rebecca

Thank you for your request for an amendment to your research proposal "Examining Measurements of Speech Disorder" as outlined in your email dated 11th July 2017.

I am pleased to advise that this request has been considered and approved by the Human Ethics Committee.

Yours sincerely

R. Robinson
pp.

Associate Professor Jane Maidment
Chair, Human Ethics Committee

Appendix B.**Speaker Demographic and Etiological Information***Table 2.*– Dysarthric Speaker Information

Speaker	Age	Aetiology	Sex	Subtype	Severity
1	53	Multiple Sclerosis	Female	Spastic-Ataxic	Mild - Moderate
2	46	Brain Tumour	Female	Ataxic	Mild-Moderate
3	47	Huntington's Disease	Female	Hyperkinetic	Moderate - Severe
4	43	Hydrocephalus	Male	Ataxic-Hypokinetic-Flaccid	Severe
5	64	Spinocerebellar Ataxia	Male	Spastic-Ataxic	Severe
6	72	Stroke	Male	Spastic-Ataxic	Severe
7	69	Cerebral Palsy	Female	Spastic	Severe
8	47	Traumatic Brain Injury	Male	Spastic-Ataxic	Severe

Note. Age taken from time of recorded sample, subtype and severity both established by consensus of speech language therapists before study onset.

Table 3. Control Speaker Information

Control Speaker	Age	Sex
1	56	Female
2	38	Female
3	49	Female
4	42	Male
5	56	Male
6	79	Male
7	68	Female
8	53	Male

Note. Age taken from time of sample recorded

Appendix C.**Listener Demographic and Pre Test Information***Table 4.* Listener Demographic and Pre Test Information

Listener	Age	PTA	MoCA
1	24	5	30
2	28	15	28
3	23	5	30
4	24	10	29
5	27	10	28
6	25	5	30
7	18	10	28
8	22	15	29
9	20	10	30
10	19	10	30
11	18	15	27
12	27	10	27
13	18	15	30
14	18	20	29
15	18	10	30
16	22	15	27
17	83	35	30
18	72	30	26
19	83	20	27
20	69	30	26
21	69	20	28
22	73	15	26
23	70	25	26
24	70	30	28
25	79	35	27
26	71	30	30
27	77	45	26
28	78	55	26
29	68	30	20
30	70	35	28
31	78	50	28
32	77	30	27

Note. Maximum score on MoCA is 30, minimum pass score is 26.

Appendix D**Sentence Intelligibility Test Sentences by Speaker**

P1

- 1) It was a great meeting
- 2) I have had my bell rung
- 3) This is a good place, though small
- 4) We selected places where families will be welcome
- 5) The first health revolution was the control of infection
- 6) They were lined with rabbit fur to protect tender feet
- 7) The driver went a short distance and made an unscheduled stop
- 8) It makes prospecting more effective and also lowers the cost of drilling
- 9) It shows you what to do if you ever want to get one
- 10) We got to talking, and I mentioned that I do candid photography of children
- 11) No one had to tell what organized labour could do for working men and women

P2

- 1) That plant needs more water
- 2) We knew we were doing well
- 3) The team had me do the throwing
- 4) After that, nature should do it for you
- 5) He is accused of being too rigid and moral
- 6) This time I just sat there and punched the pillow
- 7) You may camp in both canyons, but there are no facilities
- 8) Naturally, these new nations varied in terms of size, population, and resources
- 9) Two years ago, he heard encouraging news about a couple of old friends
- 10) I was happy to come up with a good grade out of that class
- 11) I figured on saying how much I respected him, but I didn't have a chance

P3

- 1) The job provides many benefits
- 2) The pair of shoes was new
- 3) We saw three deer and a badger
- 4) He is capable and willing to make decisions
- 5) There is nothing in this world he cannot do
- 6) Would you be asking too much to print this this day?
- 7) We don't have to worry about money we used to
- 8) For all of my long, hard years, I had been vibrantly healthy
- 9) The painters are both new fathers and love examples that I showed to them
- 10) Almost daily, the three were threatened with adverse weather conditions at the mountain cabin
- 11) That will give us a head start for the final session, which begins next month

P4

- 1) I think I'm getting better
- 2) The ballet is about to begin
- 3) Adding extra sugar didn't increase the sweetness
- 4) I soon found myself muttering to the walls
- 5) You're not supposed to be talking during the performance
- 6) My sister has a unique way of getting things done
- 7) This can be the cheapest way to ship them long distance
- 8) For years, a friend of mine has been trying to quit smoking
- 9) After a long marriage, the couple was on the verge of breaking up
- 10) I had put on thirty pounds in six months, and that didn't help matters
- 11) The inn is an unusual retreat built into the ruins on a long, black beach

P5

- 1) He took heart and played
- 2) The book is small and lightweight
- 3) Gradually, the noise level began to drop
- 4) I put on twenty pounds in one month
- 5) Distinguished law firms have volunteered to provide legal assistance
- 6) They exchanged gifts at the front door of the residence
- 7) Finding someone willing to trade with you takes a little courage
- 8) The new toaster was so shiny you could see yourself in it
- 9) When a full second is over, an addressing machine snatches the letter away
- 10) We must treat him exactly as if he were a member of our family
- 11) I untangled it, pulled him back into the boat, and got back to the dock

P6

- 1) I just couldn't feel anything
- 2) The two potatoes must be eaten
- 3) I like them because they are practical
- 4) He twisted a knee and hurt an ankle
- 5) Eventually, of course, we all got used to it
- 6) As societies become wealthier, their consumption of animal products increases
- 7) The trainees collected payment, then delivered the tickets to the customers
- 8) In any case, it is a little late to stop the revolution
- 9) After you've finished answering all the questions, please mail the card to us
- 10) Unless the young people really believe our institutions, the system will not work
- 11) He told the crowd he'd never seen anything like the celebration in his entire career

P7-

- 1) The islands are sparsely populated
- 2) Small ferns grow well in driftwood
- 3) He is definitely a notch above us
- 4) The defendant is either guilty or not guilty
- 5) She has never met a man like me before
- 6) Things obviously were less tense there than she had pictured
- 7) The ice cream you wanted for a midnight snack has vanished
- 8) After what seemed like hours of waiting, the taxi finally showed up
- 9) There was no history of the disease on either side of my family
- 10) If this doesn't work, he should be mature enough to get good professional help
- 11) In order to win, we must put more points on the board than we have

P8

- 1) The sun died at night
- 2) Sunlight shone through the bedroom window
- 3) I gave the role everything I had
- 4) That's what I thought it was at first
- 5) I can only describe that action as totally ridiculous
- 6) Most studies of animal behaviour do not support this view
- 7) The basic requirement is a stimulating activity that draws them closer
- 8) A cat raised by other cats will be forever fearful of people
- 9) Next evening, we were back to take our places against the fallen tree
- 10) Their brakes are much more reliable and their cooling systems require much less maintenance
- 11) Due to amnesia, several men would remember nothing of the crash from that point on.

Appendix E**Veracity Statements and Ambiguity Ratings***Table 5.* Veracity Statements and Ambiguity Ratings

Speaker	Sentence	Ambiguity Rating
P1		
	Brooms clean floors	5
	Red is blue	5
	Computers have keyboards	5
	Ovens sew blankets	5
	Fire is hot	5
	Polkadots are stripes	5
	Plants can grow	5
	Rabbits eat steak	5
	Snakes are reptiles	5
	Cheese is purple	5
	Cash is money	5
	Noses can hear	5
	Films are movies	5
	Forests have beards	5
	Babies wear nappies	5
	Princes are women	5
	Jails have bars	5
	Monkeys are produce	5
	Cows chew grass	5
	Fingers get sick	5
P2		
	Frogs wear clothes	5
	Shoes cover feet	5
	Benches have brains	5
	Sugar is sweet	5
	Cars use statues	5

Bombs can explode	5
Dogs play pianos	5
Elephants have ears	5
Pencils take pictures	5
Mittens cover hands	5
Yarn is hard	5
Islands have beaches	5
Houses have engines	5
Tongues tastes flavour	5
Chips are sweet	4
Carrots are healthy	5
Water is dry	5
Students read books	4
Pumpkins are tables	5
Carrot are oranges	5
Snails are fruit	5
Cowboys ride horses	5
Snails are fast	4
Knives can cut	5
Skateboards can fight	5
Farmers grow cabbage	4
Sandpaper is smooth	4
Yards are outside	5
Soap is dirty	5
Theatres play movies	5
Chairs have conversations	5
Planes can fly	5
Giraffes are short	4
Birds have wings	5
Glasses help vision	5
Concerts are quiet	4
Balls can roll	5

P3

P4	Leopards have fingers	4
	Giraffes have spots	4
<hr/>	Swimming is exercise	5
	Squares are circles	5
	Families take photos	4
	Marshmallows are hard	5
	Notebooks are paper	4
	Zebras have spots	5
	Lemons are sour	5
	Purses are elephants	5
	Rockets can fly	5
	Kittens are puppies	5
	Illustrators draw pictures	5
	Onions wear glasses	5
	Spiders create webs	5
	Hammers rake leaves	5
	Bananas are fruit	5
	Pigs quilt blankets	5
	Chickens have beaks	5
	Bombs are safe	5
	Oranges are nutritious	5
	Heat is cold	5
<hr/>	Thorns are dull	4
	Applications are reviewed	-
	Clothes heat food	5
	Cameras take pictures	5
	Elephants wear shoes	5
	Carbohydrates are fuel	3
	Turtles are fast	5
	Cake is a dessert	4
	Lizards wear rings	-

Radios play music	5
Cacti grow apples	5
Mail needs stamps	5
Queens are men	4
Guitars make music	5
Computers make rocks	5
Penguins can waddle	5
Purses hold rice	5
Phones make calls	5
Rubbish smells nice	5
Frogs eat flies	5

P6

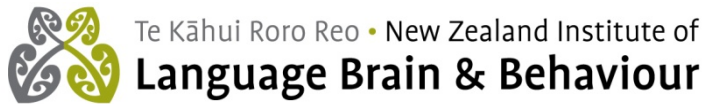
Cars have headlights	5
Mints are salty	5
Coal is black	5
Pianos play chess	5
Feathers are soft	5
Ears can taste	5
Spring brings flowers	5
Snakes can sing	5
Frogs take leaps	5
Concrete is soft	5
Butterflies can flutter	5
Sandals are dresses	5
Policemen fight crime	5
Straws are hair	5
Earth is round	5
Scissors cut rocks	5
Rainbows are colourful	5
Pets make jewellery	5
Frogs catch bugs	4
Muffins are fruits	5

P7

	Seeds grow hair	4
	Bikes have handlebars	5
	Teaches lecture mice	5
	Pens need ink	5
	Wallets hold water	5
	Horses eat hay	4
	People hate laughing	4
	People watch television	5
	Mountains are little	5
	Turtles are slow	5
	Rabbits jump rope	4
	Restaurants serve food	5
	Trees have fingers	5
	Spiders have legs	5
	Insects are mammals	5
	Noses smell scents	5
	Cats fly high	5
	Toddlers can play	5
	Eyes can hear	5
	Hospitals have doctors	5
P8		
	Mice are fast	4
	Bikes have doors	5
	Mirrors can reflect	5
	Juices drink people	5
	Tyedye is colourful	4
	Notebooks have feelings	5
	Beaches have sand	5
	Socks are shapes	-
	Circles are shapes	5
	Babies often drive	5
	Trees produce oxygen	5
	Frisbees are balls	5

Flamingos are pink	4
Parades are silent	-
Water rings loud	-
Wallets carry money	4
Men eat pencils	5
Farmers sow seeds	3
People shout quietly	-
Kangaroos can jump	-

note. Veracity statements read but excluded are in bold. If no rating next to word it was excluded before ratings were made for reasons other than ambiguity.

Appendix F**Study Information Sheet for Listeners**

Project Name: Examining listener reaction time as a measure of speech disorder

You are invited to take part in the research project titled: Reaction time as a measure of listener perception. Please take the time to read this information sheet thoroughly and consider whether you would like to participate. Your participation is entirely voluntary (your choice). The following research team is conducting this study:

Rebecca Risi
Master of Audiology Thesis Student
rebecca.risi@pg.canterbury.ac.nz
0273501900

Dr Annalise Fletcher
Lecturer, Department of Communication Disorders & NZILBB
annalise.fletcher@canterbury.ac.nz
+64 3 3695691, Ext 95691

Dr Megan McAuliffe
Associate Professor, Department of Communication Disorders
Theme Leader, Language & Ageing, NZILBB

STUDY INFORMATION

This study will look at the time taken for healthy New Zealanders aged between 18- 30 and 65-90 years to react to speech from a speaker with dysarthria (a disorder affecting how people speak). It will examine how well this measure compares to both rating scales and transcription of speech samples to see if this can be used as a measure of severity of a speech disorder in adults. It will also determine if there are difference in the way younger and older listeners react to disordered speech.

To carry out this research we are looking for 30 people from two age groups—those aged 18 to 30 years, and those aged 65 to 90 years. If you agree to be involved, we will ask you to complete a series of assessments at our research laboratory at the University of Canterbury. Your participation will involve two sections, and both will last approximately 30 to 45 minutes. To be involved in this study, you should fit the age criteria and: (1) be a native speaker of New Zealand English, familiar listening to NZ speakers, (2) have no previous history of learning disorder (3) have no previous history of significant memory or cognitive problems, and (4) consider your hearing to be age-appropriate.

If you choose to undertake this project you will be asked to complete the following tasks, in a random order at our research laboratory:

Listening experiment 1: You will hear short phrases and be asked to say what you think you hear. Some of these will be easy to understand, some may be very difficult. You will simply be asked to say what you think you hear into a microphone, and this will be recorded. These results will then be analysed to determine how much of what was said you actually understood, and what kind of errors you made. (approximately 10 minutes).

Listening experiment 2: This is similar to experiment 1, except instead of saying what you heard, you will rate the speaker from 1-10 in terms of their intelligibility (approximately 10 minutes)

Listening Experiment 3: You will hear a short statement that is true or false. You will listen and then select either true or false on the laptop. Some of these will be easy to understand, some may be very difficult, but it is important to select which one you think is correct (approximately 10 minutes).

Hearing test: This will take about 10 minutes, and during this time you will be asked to respond when you hear a “beep” noise. If we notice you do have some difficulty hearing (which is expected for many participants in the older age group), you may also wish to have a full hearing assessment at the Speech and Hearing Clinic, Department of Communication Disorders, University of Canterbury. We can organise this for you if you would like to follow up.

Cognition assessment: you will be asked to complete a series of simple tasks that include examining your ability to name items, match patterns and recall words and digits, do simple maths problems and follow instructions. If you find this task very difficult we will discontinue testing as suggest that you visit your GP for further assessment.

When you have finished the experiments, we will discuss the study with you and answer any questions you might have about your performance. We can send you a summary of the study findings when it has been completed (this should be completed by February 2018).

CONFIDENTIALITY

Your privacy and confidentiality will be maintained at all times. Your completed consent forms will be kept in a locked filing cabinet within a lockable room at our research laboratory, or within the Department of Communication Disorders. Similarly, papers from your hearing, cognition and linguistic knowledge assessments will be identified by a numeric code only and stored in a locked filing cabinet, with scanned digitized copies kept in our lab’s secured access folder. Your voice recordings taken during the perception experimentation (i.e., when you say what you think you are hearing), will also be stored via a numeric code only, and the recordings kept on our secure access folder. With your consent, your data (paper and voice-based) will be stored for use in this and future projects that have received ethical clearance from the UC Human Ethics Committee – and in teaching and research presentations associated with this research.

BENEFIT

Participating in this research will have no direct benefits for you personally. However, we expect that findings from this study will inform researchers on the reliability of different methods of assessing speech disorders, which may lead to the development of more objective

assessment tools, ideally improving quality of life for tens of thousands of New Zealanders and potentially hundreds of millions worldwide.

REIMBURSEMENT FOR PARTICIPATION

On completion of the assessment, you will receive a 3 course credits towards CMDS 162/667 as reimbursement for your time. If you are not enrolled in this course you will instead receive a \$20 voucher.

RISKS OF PARTICIPATION

This research study poses no risks to personal health. If we detect some degree of hearing loss on the brief hearing assessment, with your consent we will refer you to the hearing clinic at the University of Canterbury for a full assessment.

WITHDRAWING FROM THE STUDY

It is important to note that this study is voluntary and that you can withdraw from it at any time. This will in no way jeopardise any of your future dealings with the Department of Communication Disorders or New Zealand Institute of Language, Brain and Behaviour. If you choose to withdraw from the study, any data collected prior to withdrawal will not be used for research purposes without your consent.

ETHICS

The University of Canterbury Human Ethics Committee has approved this project. Please do not hesitate to contact Rebecca Risi if you have any concerns regarding your participation in this project. Alternatively, if there are any concerns you may contact Dr Annalise Fletcher at annalise.fletcher@pg.canterbury.ac.nz.

FOR MORE INFORMATION

Should you have further questions regarding the research, please feel free to contact Rebecca Risi (rebecca.risi@pg.canterbury.ac.nz)

THANK YOU FOR YOUR INTEREST IN THIS RESEARCH

Appendix G

Consent Form for Listeners



CONSENT FORM: LISTENERS

Project Name: Examining Listener reaction time as a measure of speech disorder

Investigators: Rebecca Risi (MAud student)

Dr Annalise Fletcher
Lecturer, Department of Communication
Disorders & NZILBB

Dr Megan McAuliffe
Professor, Department of Communication Disorders
Theme Leader, Language & Ageing, NZILBB

Please read the following carefully:

- I have read and I understand the information sheet for volunteers taking part in the study designed to evaluate different measures of rating disordered speech.
- I have had the opportunity to discuss this study with the researcher/s. I am satisfied with the answers I have been given.
- I understand that my participation in this study is confidential and that no material that could identify me will be used in any reports on this study.
- I have had time to consider whether to take part.
- I understand that taking part in this study is voluntary (my choice) and that I may withdraw from the study at any time and this will in no way affect my future health care.

I consent to my data collected being made available for future studies that have received ethical clearance from the University of Canterbury Ethics Committee.

YES/NO

I wish to receive a copy of the results.

YES/NO

I hereby consent to take part in the study:

NAME (please print): _____

Signature: _____ Date: _____

Project explained by: _____

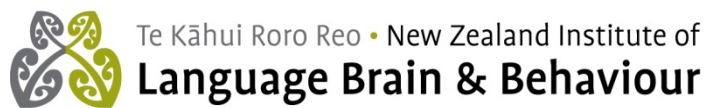
Project role: _____

Signature: _____

Date: _____

Appendix H

Demographic Questionnaire for listeners



Demographic Questionnaire

The information collected below will not be published, and you will not be identified at any point in this study using this information. It will be kept separate from your results during the assessments.

NAME (please print): _____

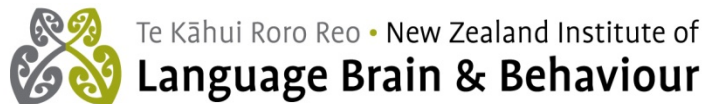
Date of Birth: _____ Age: _____

Address: _____ (to send copy of results if required)

Contact number: _____

Highest level of Education: _____

Employment/Previous Occupation: _____

Appendix I**Information for Listeners Requiring Full Diagnostic Hearing Test**

Dear:

Re: Hearing screening assessment as part of the research project “Examining Listener Reaction Times as a Measure of Speech Disorder”.

Thank you for participating in the research project “**Examining Listener Reaction Times as a Measure of Speech Disorder**”. As part of this research, you completed a pure tone hearing assessment. This is a quick screening assessment of hearing. A screening assessment is used to identify whether people *might* have a problem in a certain area. It does not determine the nature or severity of the problem.

Your score on the assessment was outside of what we would generally expect, which suggests there might be a hearing loss. We suggest that you attend a full diagnostic hearing test. With your consent, we will forward a copy of this letter and your hearing screen to the University of Canterbury Speech and Hearing Clinic so that you may discuss this further.

If you require further clarification, please do not hesitate to contact me. Thank you again for participating in our research.

Regards,

Rebecca Risi

Master of Audiology Student

rebecca.risi@pg.canterbury.ac.nz

0273501900

Annalise Fletcher PhD

Supervisor

Department of Communication Disorders

New Zealand Institute of Language, Brain & Behaviour

Email: annalise.fletcher@canterbury.ac.nz

+64 3 3695691, Ext 95691

Appendix J**Montreal Cognitive Assessment, and consent to use on behalf of Dr Ziad Nasreddine****MoCa**

Wednesday, 26 July 2017 at 3:24 AM

To: risi.rebecca8@gmail.com

Cc: Ziad Nasreddine

Hello,

Thank you for your interest in the MoCA®.

You are welcome to use the MoCA® Test as you described below with no further permission requirements.

No changes or adaptations to the MoCA® Test and instructions are permitted.

All the best,

**Kathleen Gallant, MSOT**

Occupational Therapist/ Psychometrician

On behalf of Dr Ziad Nasreddine, Neurologist, MoCA® Copyright Owner

MoCA Clinic & Institute

4896 Taschereau Blvd, suite 230

Greenfield Park, Quebec, Canada, J4V 2J2

Tel : (450) 672-7766 #222 Fax : (450) 672-3899

kathleen.gallant@mocaclinic.cawww.mocatest.org / www.alzheimer.TV

MONTREAL COGNITIVE ASSESSMENT (MOCA)

NAME :

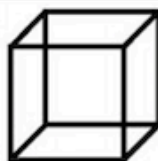
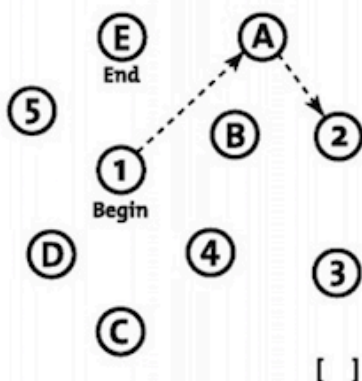
Education :

Sex :

Date of birth :

DATE :

VISUOSPATIAL / EXECUTIVE

Copy
cubeDraw CLOCK (Ten past eleven)
(3 points)

POINTS

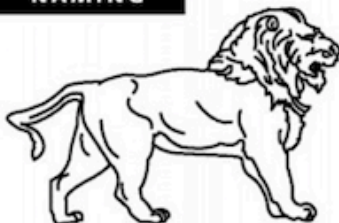
[]

[]

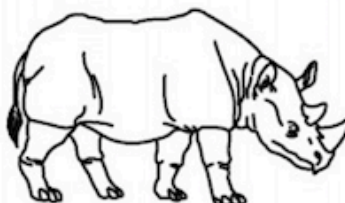
[]
Contour[]
Numbers[]
Hands

___/5

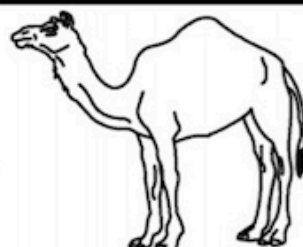
NAMING



[]



[]



[]

___/3

MEMORY

Read list of words, subject
must repeat them. Do 2 trials.
Do a recall after 5 minutes.

	FACE	VELVET	CHURCH	DAISY	RED
1st trial					
2nd trial					

No
points

ATTENTION

Read list of digits (1 digit/ sec.).

Subject has to repeat them in the forward order

[] 2 1 8 5 4

Subject has to repeat them in the backward order

[] 7 4 2

___/2

Read list of letters. The subject must tap with his hand at each letter A. No points if ≥ 2 errors

[] FBACMNAAJKLBAFAKDEAAAJAMOF AAB

___/1

Serial 7 subtraction starting at 100

[] 93

[] 86

[] 79

[] 72

[] 65

4 or 5 correct subtractions: 3 pts, 2 or 3 correct: 2 pts, 1 correct: 1 pt, 0 correct: 0 pt

___/3

LANGUAGE

Repeat : I only know that John is the one to help today. []

The cat always hid under the couch when dogs were in the room. []

___/2

Fluency / Name maximum number of words in one minute that begin with the letter F [] _____ (N ≥ 11 words)

___/1

ABSTRACTION

Similarity between e.g. banana - orange = fruit [] train - bicycle [] watch - ruler

___/2

DELAYED RECALL

Has to recall words

WITH NO CUE

FACE

[]

VELVET

[]

CHURCH

[]

DAISY

[]

RED

[]

Points for
UNCUED
recall only

___/5

Optional

Category cue

Multiple choice cue

[]

[]

[]

[]

[]

ORIENTATION

[] Date

[] Month

[] Year

[] Day

[] Place

[] City

___/6